

Gencore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 12:03:31 : Search time 12 Seconds  
(without alignments)  
519.754 Million cell updates/sec

Title: US-09-786-926-1

Perfect score: 1966

Sequence: 1 NNATCTPVAPESCQQLAAG.....RSLGFRMRKPLSTISYSRSI 384

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCIT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCITUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 1966  | 100.0       | 384    | 9  | US-09-971-228-10  |
| 2          | 1966  | 100.0       | 384    | 10 | US-09-842-316-9   |
| 3          | 1966  | 100.0       | 384    | 10 | US-09-731-030A-11 |
| 4          | 1966  | 100.0       | 384    | 12 | US-10-007-399-1   |
| 5          | 1966  | 100.0       | 384    | 12 | US-10-037-616-24  |
| 6          | 1956  | 99.5        | 384    | 10 | US-09-731-030A-13 |
| 7          | 1948  | 99.1        | 384    | 10 | US-09-827-937A-4  |
| 8          | 1642  | 83.5        | 386    | 10 | US-09-917-508-2   |
| 9          | 993   | 50.5        | 224    | 10 | US-09-731-030A-21 |
| 10         | 653   | 33.2        | 382    | 10 | US-09-969-711-2   |
| 11         | 645.5 | 32.8        | 378    | 9  | US-09-971-228-7   |
| 12         | 645.5 | 32.8        | 378    | 10 | US-09-842-316-7   |
| 13         | 645.5 | 32.8        | 378    | 10 | US-09-731-030A-18 |
| 14         | 645.5 | 32.8        | 378    | 12 | US-10-037-616-21  |
| 15         | 645.5 | 32.8        | 398    | 10 | US-09-815-333-2   |
| 16         | 645.5 | 32.8        | 398    | 10 | US-09-842-316-2   |
| 17         | 645.5 | 32.8        | 398    | 12 | US-10-037-616-26  |
| 18         | 640.5 | 32.6        | 362    | 10 | US-09-993-844-8   |
| 19         | 638   | 32.5        | 400    | 9  | US-09-971-228-12  |

|    |       |      |     |    |                   |                   |
|----|-------|------|-----|----|-------------------|-------------------|
| 20 | 638   | 32.5 | 400 | 10 | US-09-812-272-2   | Sequence 2, Appli |
| 21 | 638   | 32.5 | 400 | 12 | US-10-037-616-27  | Sequence 27, Appl |
| 22 | 635.5 | 32.3 | 341 | 9  | US-09-971-228-13  | Sequence 13, Appl |
| 23 | 635.5 | 32.3 | 381 | 12 | US-10-037-616-19  | Sequence 19, Appl |
| 24 | 635.5 | 32.3 | 381 | 10 | US-09-971-228-5   | Sequence 5, Appli |
| 25 | 633   | 32.2 | 381 | 10 | US-09-827-937A-18 | Sequence 18, Appl |
| 26 | 633   | 32.2 | 381 | 10 | US-09-842-316-6   | Sequence 6, Appli |
| 27 | 633   | 32.2 | 381 | 10 | US-09-731-030A-17 | Sequence 17, Appl |
| 28 | 627.5 | 31.9 | 346 | 9  | US-09-971-228-15  | Sequence 15, Appl |
| 29 | 605   | 30.8 | 314 | 9  | US-09-971-228-14  | Sequence 14, Appl |
| 30 | 535   | 27.2 | 353 | 10 | US-09-731-030A-19 | Sequence 19, Appl |
| 31 | 534   | 27.2 | 353 | 10 | US-09-771-063-2   | Sequence 2, Appli |
| 32 | 534   | 27.2 | 353 | 12 | US-10-037-616-23  | Sequence 23, Appl |
| 33 | 526   | 26.8 | 353 | 9  | US-09-971-228-9   | Sequence 9, Appli |
| 34 | 526   | 26.8 | 353 | 10 | US-09-842-316-8   | Sequence 8, Appli |
| 35 | 520   | 26.4 | 353 | 10 | US-09-771-063-4   | Sequence 4, Appli |
| 36 | 479.5 | 24.4 | 351 | 9  | US-09-971-228-8   | Sequence 8, Appli |
| 37 | 479.5 | 24.4 | 351 | 12 | US-10-037-616-22  | Sequence 22, Appl |
| 38 | 479.5 | 24.4 | 382 | 10 | US-09-842-316-5   | Sequence 5, Appli |
| 39 | 450.5 | 22.9 | 348 | 10 | US-09-903-799-2   | Sequence 2, Appli |
| 40 | 438   | 22.3 | 364 | 9  | US-09-971-228-6   | Sequence 6, Appli |
| 41 | 438   | 22.3 | 364 | 10 | US-09-842-316-3   | Sequence 3, Appli |
| 42 | 438   | 22.3 | 364 | 10 | US-09-731-030A-15 | Sequence 15, Appl |
| 43 | 438   | 22.3 | 364 | 12 | US-10-037-616-20  | Sequence 20, Appl |
| 44 | 425.5 | 21.6 | 354 | 10 | US-09-879-225-2   | Sequence 2, Appli |
| 45 | 416.5 | 21.2 | 353 | 9  | US-09-971-228-11  | Sequence 11, Appl |

#### ALIGNMENTS

##### RESULT 1

US-09-971-228-10  
Sequence 10, Application US/09971228  
Patent No. US2002015512A1  
GENERAL INFORMATION:  
APPLICANT: Liao, X, Charlene  
APPLICANT: Masuda, Esteban  
APPLICANT: Chu, Peter  
APPLICANT: Pardo, Jorge  
APPLICANT: Li, Congfen  
APPLICANT: Zhao, Haoran  
APPLICANT: Jiang, Yiluying  
APPLICANT: Rigel Pharmaceuticals, Incorporated  
TITLE OF INVENTION: EDG: Modulators of Lymphocyte Activation and Migration  
FILE REFERENCE: 021044-000310US  
CURRENT APPLICATION NUMBER: US/09/971,228  
CURRENT FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: US 60/284,763  
PRIOR FILING DATE: 2001-04-18  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 10  
LENGTH: 384  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: human endothelial differentiation G-protein  
OTHER INFORMATION: coupled receptor (GPCR) 6 (RDG6)  
US-09-971-228-10

Query Match 100.0%; Score 1966; DB 9; Length 384;  
Best Local Similarity 100.0%; Pred. No. 5e-159;  
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNATCTPVAPESCQQLAAGSHRLIVLHYNHSRGLAGGEGDGLGALRGLSYAASCIV 60  
DB 1 NNATCTPVAPESCQQLAAGSHRLIVLHYNHSRGLAGGEGDGLGALRGLSYAASCIV 60  
QY 61 VLENLIVLAATISHRSRRWYCYLVNTITSDLTGAIVLANVILSGARFRLPAQWFL 120  
DB 61 VLENLIVLAATISHRSRRWYCYLVNTITSDLTGAIVLANVILSGARFRLPAQWFL 120

QY 121 REGLLFTALAASTFSLFTAGERFATWVPVAESGATKTSRYVFGICGLCWLALALGMLP 180  
Db 121 REGLLFTALAASTFSLFTAGERFATWVPVAESGATKTSRYVFGICGLCWLALALGMLP 180  
QY 181 LLGNMCLCAFDRCSLLPLYSKRYILFCLVIFAGVLAITMGLYGAIFRLVQASGOKAPRP 240  
Db 181 LLGNMCLCAFDRCSLLPLYSKRYILFCLVIFAGVLAITMGLYGAIFRLVQASGOKAPRP 240  
QY 241 AARRKARRLKTVMLILAFVCMGPLEFGLLADVFGSNLMAOEYLRGMWIIALAVLNS 300  
Db 241 AARRKARRLKTVMLILAFVCMGPLEFGLLADVFGSNLMAOEYLRGMWIIALAVLNS 300  
QY 301 AVNDIYSFRSREVCRAVLSFLCCGCLRLGMRGPDCLARAVERAHSASTTDSLSLRPDS 360  
Db 301 AVNDIYSFRSREVCRAVLSFLCCGCLRLGMRGPDCLARAVERAHSASTTDSLSLRPDS 360  
QY 361 FRGSRLSFRMRREPLSSISVSRSI 384  
Db 361 FRGSRLSFRMRREPLSSISVSRSI 384

## RESULT 2

US-09-842-316-9  
; Sequence 9, Application US/09842316  
; Patent No. US2002009191A1  
; GENERAL INFORMATION:  
; APPLICANT: KOSTENSIS, Eva  
; APPLICANT: GASENHUBER, Johann  
; TITLE OF INVENTION: EDG8 RECEPTOR, ITS PREPARATION AND USE  
; FILE REFERENCE: 38005-147  
; CURRENT APPLICATION NUMBER: US/09/842,316  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: EP 116589.3  
; PRIOR FILING DATE: 2000-08-01  
; PRIOR APPLICATION NUMBER: EP 108658.2  
; PRIOR FILING DATE: 2000-04-26  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-842-316-9

Query Match 100.0%; Score 1966; DB 10; Length 384;  
Best Local Similarity 100.0%; Pred. No. 5e-159;

Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNATGTPVAPESCQQLAAGHSRLIVLHYNHSGRLAGRGPEDEGLGALRGLSVAASCLV 60  
Db 1 MNATGTPVAPESCQQLAAGHSRLIVLHYNHSGRLAGRGPEDEGLGALRGLSVAASCLV 60  
QY 61 VLENLVLVAITSHMRSRRWYVYCLVNITLSDLTGAAYLANVLGSGARTFRLAPQWFL 120  
Db 61 VLENLVLVAITSHMRSRRWYVYCLVNITLSDLTGAAYLANVLGSGARTFRLAPQWFL 120  
QY 121 REGLLFTALAASTFSLFTAGERFATWVPVAESGATKTSRYVFGICGLCWLALALGMLP 180  
Db 121 REGLLFTALAASTFSLFTAGERFATWVPVAESGATKTSRYVFGICGLCWLALALGMLP 180  
QY 181 LLGNMCLCAFDRCSLLPLYSKRYILFCLVIFAGVLAITMGLYGAIFRLVQASGOKAPRP 240  
Db 181 LLGNMCLCAFDRCSLLPLYSKRYILFCLVIFAGVLAITMGLYGAIFRLVQASGOKAPRP 240  
QY 241 AARRKARRLKTVMLILAFVCMGPLEFGLLADVFGSNLMAOEYLRGMWIIALAVLNS 300  
Db 241 AARRKARRLKTVMLILAFVCMGPLEFGLLADVFGSNLMAOEYLRGMWIIALAVLNS 300  
QY 301 AVNDIYSFRSREVCRAVLSFLCCGCLRLGMRGPDCLARAVERAHSASTTDSLSLRPDS 360  
Db 301 AVNDIYSFRSREVCRAVLSFLCCGCLRLGMRGPDCLARAVERAHSASTTDSLSLRPDS 360  
QY 361 FRGSRLSFRMRREPLSSISVSRSI 384

Db 361 FRGSRLSFRMRREPLSSISVSRSI 384

## RESULT 3

US-09-731-030A-11  
; Sequence 11, Application US/09731030A  
; Patent No. US20020142375A1  
; GENERAL INFORMATION:  
; APPLICANT: MONROE, Donald G.  
; APPLICANT: GUPTA, Ashwani K.  
; APPLICANT: ZASTAWNY, Roman L.  
; TITLE OF INVENTION: MAMMALIAN EDG-7 RECEPTOR HOMOLOGS  
; FILE REFERENCE: 8074-0015  
; CURRENT APPLICATION NUMBER: US/09/731,030A  
; PRIOR FILING DATE: 1996-12-29  
; PRIOR APPLICATION NUMBER: 60/070,184  
; PRIOR FILING DATE: 1997-12-30  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-731-030A-11

Query Match 100.0%; Score 1966; DB 10; Length 384;  
Best Local Similarity 100.0%; Pred. No. 5e-159;

Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNATGTPVAPESCQQLAAGHSRLIVLHYNHSGRLAGRGPEDEGLGALRGLSVAASCLV 60  
Db 1 MNATGTPVAPESCQQLAAGHSRLIVLHYNHSGRLAGRGPEDEGLGALRGLSVAASCLV 60  
QY 61 VLENLVLVAITSHMRSRRWYVYCLVNITLSDLTGAAYLANVLGSGARTFRLAPQWFL 120  
Db 61 VLENLVLVAITSHMRSRRWYVYCLVNITLSDLTGAAYLANVLGSGARTFRLAPQWFL 120  
QY 121 REGLLFTALAASTFSLFTAGERFATWVPVAESGATKTSRYVFGICGLCWLALALGMLP 180  
Db 121 REGLLFTALAASTFSLFTAGERFATWVPVAESGATKTSRYVFGICGLCWLALALGMLP 180  
QY 181 LLGNMCLCAFDRCSLLPLYSKRYILFCLVIFAGVLAITMGLYGAIFRLVQASGOKAPRP 240  
Db 181 LLGNMCLCAFDRCSLLPLYSKRYILFCLVIFAGVLAITMGLYGAIFRLVQASGOKAPRP 240  
QY 241 AARRKARRLKTVMLILAFVCMGPLEFGLLADVFGSNLMAOEYLRGMWIIALAVLNS 300  
Db 241 AARRKARRLKTVMLILAFVCMGPLEFGLLADVFGSNLMAOEYLRGMWIIALAVLNS 300  
QY 301 AVNDIYSFRSREVCRAVLSFLCCGCLRLGMRGPDCLARAVERAHSASTTDSLSLRPDS 360  
Db 301 AVNDIYSFRSREVCRAVLSFLCCGCLRLGMRGPDCLARAVERAHSASTTDSLSLRPDS 360  
QY 361 FRGSRLSFRMRREPLSSISVSRSI 384  
Db 361 FRGSRLSFRMRREPLSSISVSRSI 384

## RESULT 4

US-10-007-399-1  
; Sequence 1, Application US/10007399  
; Patent No. US20020115150A1  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria A.  
; TITLE OF INVENTION: 14275 Receptor, A No. US20020115150A1el G-Protein Coupled Receptor Family  
; FILE REFERENCE: 035800/169287  
; CURRENT APPLICATION NUMBER: US/10/007,399  
; PRIOR FILING DATE: 2001-11-05  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1

LENGTH: 384  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-007-399-1

Query Match 100.0%; Score 1966; DB 12; Length 384;  
Best Local Similarity 100.0%; Pred. No. 5e-159;  
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNATGTPVAPESCOOLAGGHSRLIVLHYNHSGLAGRGPEDEGGIGALRGLSVAASCLV 60  
DB 1 MNATGTPVAPESCOOLAGGHSRLIVLHYNHSGLAGRGPEDEGGIGALRGLSVAASCLV 60  
QY 61 VLENILVLAATITSHMRSRRWYVCLVNTITSDLTGAAVLAVLNSGARTFRLAPQWFL 120  
DB 61 VLENILVLAATITSHMRSRRWYVCLVNTITSDLTGAAVLAVLNSGARTFRLAPQWFL 120  
QY 121 REGILFTALAASTFSLFTAGERFATWVRPAESGATKTSRYVGFGLCWLALALGMLP 180  
DB 121 REGILFTALAASTFSLFTAGERFATWVRPAESGATKTSRYVGFGLCWLALALGMLP 180  
QY 181 ILGWNCLAFDRCSLLPLYSKRYILFCLVIFAGVLATIMGLYGAIFRLVQASGOKARPP 240  
DB 181 ILGWNCLAFDRCSLLPLYSKRYILFCLVIFAGVLATIMGLYGAIFRLVQASGOKARPP 240  
QY 241 AARRKARRLKTYMLLAFVCMGPLEGILLADVFSGNMAOEYLRGMDWILALAVLNS 300  
DB 241 AARRKARRLKTYMLLAFVCMGPLEGILLADVFSGNMAOEYLRGMDWILALAVLNS 300  
QY 301 AVNPITYSFRSREVCRAVLSFLCCGCLRLGMRGPGDCLARAWEAHSASTTDSLRPRS 360  
DB 301 AVNPITYSFRSREVCRAVLSFLCCGCLRLGMRGPGDCLARAWEAHSASTTDSLRPRS 360  
QY 361 FRGSRSLSPMRREPLSSISVSRSI 384  
DB 361 FRGSRSLSPMRREPLSSISVSRSI 384

RESULT 5  
US-10-037-616-24

; Sequence 24; Application US/10037616

; Patent No. US20020123148A1

; GENERAL INFORMATION:

; APPLICANT: English, Denis

; APPLICANT: Kovacs, Richard J.

; APPLICANT: Rizzo, Maria T.

; APPLICANT: Sliva, Daniel T.

; TITLE OF INVENTION: Sphingolipid Compositions and Methods for Their Therapeutic Use

; FILE REFERENCE: 7042-119

; CURRENT APPLICATION NUMBER: US/10/037, 616

; CURRENT FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: US 60/243,887

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 24

; LENGTH: 384

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-037-616-24

Query Match 100.0%; Score 1966; DB 12; Length 384;  
Best Local Similarity 100.0%; Pred. No. 5e-159;  
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNATGTPVAPESCOOLAGGHSRLIVLHYNHSGLAGRGPEDEGGIGALRGLSVAASCLV 60  
DB 1 MNATGTPVAPESCOOLAGGHSRLIVLHYNHSGLAGRGPEDEGGIGALRGLSVAASCLV 60  
QY 61 VLENILVLAATITSHMRSRRWYVCLVNTITSDLTGAAVLAVLNSGARTFRLAPQWFL 120  
DB 61 VLENILVLAATITSHMRSRRWYVCLVNTITSDLTGAAVLAVLNSGARTFRLAPQWFL 120

QY 121 REGILFTALAASTFSLFTAGERFATWVRPAESGATKTSRYVGFGLCWLALALGMLP 180  
DB 121 REGILFTALAASTFSLFTAGERFATWVRPAESGATKTSRYVGFGLCWLALALGMLP 180  
QY 181 ILGWNCLAFDRCSLLPLYSKRYILFCLVIFAGVLATIMGLYGAIFRLVQASGOKARPP 240  
DB 181 ILGWNCLAFDRCSLLPLYSKRYILFCLVIFAGVLATIMGLYGAIFRLVQASGOKARPP 240  
QY 241 AARRKARRLKTYMLLAFVCMGPLEGILLADVFSGNMAOEYLRGMDWILALAVLNS 300  
DB 241 AARRKARRLKTYMLLAFVCMGPLEGILLADVFSGNMAOEYLRGMDWILALAVLNS 300  
QY 301 AVNPITYSFRSREVCRAVLSFLCCGCLRLGMRGPGDCLARAWEAHSASTTDSLRPRS 360  
DB 301 AVNPITYSFRSREVCRAVLSFLCCGCLRLGMRGPGDCLARAWEAHSASTTDSLRPRS 360  
QY 361 FRGSRSLSPMRREPLSSISVSRSI 384  
DB 361 FRGSRSLSPMRREPLSSISVSRSI 384

RESULT 6

US-09-731-030A-13

; Sequence 13; Application US/09731030A

; Patent No. US20020142375A1

; GENERAL INFORMATION:

; APPLICANT: MONROE, Donald G.

; APPLICANT: GUPTA, Ashwani K.

; APPLICANT: ZASTAWNY, Roman L.

; TITLE OF INVENTION: MAMMALIAN EDG-7 RECEPTOR HOMOLOGS

; FILE REFERENCE: 8074-0015

; CURRENT APPLICATION NUMBER: US/09/731,030A

; CURRENT FILING DATE: 1998-12-29

; PRIOR APPLICATION NUMBER: 60/070,184

; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 13

; LENGTH: 384

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-731-030A-13

Query Match 99.5%; Score 1956; DB 10; Length 384;  
Best Local Similarity 99.5%; Pred. No. 3.5e-158;  
Matches 382; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNATGTPVAPESCOOLAGGHSRLIVLHYNHSGLAGRGPEDEGGIGALRGLSVAASCLV 60  
DB 1 MNATGTPVAPESCOOLAGGHSRLIVLHYNHSGLAGRGPEDEGGIGALRGLSVAASCLV 60  
QY 61 VLENILVLAATITSHMRSRRWYVCLVNTITSDLTGAAVLAVLNSGARTFRLAPQWFL 120  
DB 61 VLENILVLAATITSHMRSRRWYVCLVNTITSDLTGAAVLAVLNSGARTFRLAPQWFL 120  
QY 121 REGILFTALAASTFSLFTAGERFATWVRPAESGATKTSRYVGFGLCWLALALGMLP 180  
DB 121 REGILFTALAASTFSLFTAGERFATWVRPAESGATKTSRYVGFGLCWLALALGMLP 180  
QY 181 ILGWNCLAFDRCSLLPLYSKRYILFCLVIFAGVLATIMGLYGAIFRLVQASGOKARPP 240  
DB 181 ILGWNCLAFDRCSLLPLYSKRYILFCLVIFAGVLATIMGLYGAIFRLVQASGOKARPP 240  
QY 241 AARRKARRLKTYMLLAFVCMGPLEGILLADVFSGNMAOEYLRGMDWILALAVLNS 300  
DB 241 AARRKARRLKTYMLLAFVCMGPLEGILLADVFSGNMAOEYLRGMDWILALAVLNS 300  
QY 301 AVNPITYSFRSREVCRAVLSFLCCGCLRLGMRGPGDCLARAWEAHSASTTDSLRPRS 360  
DB 301 AVNPITYSFRSREVCRAVLSFLCCGCLRLGMRGPGDCLARAWEAHSASTTDSLRPRS 360  
QY 361 FRGSRSLSPMRREPLSSISVSRSI 384  
DB 361 FRGSRSLSPMRREPLSSISVSRSI 384

Db 361 FRGSRLSFRMRPLSSSSSVRSI 384

## RESULT 7

US-09-827-937A-4  
Sequence 4, Application US/09827937A  
Patent No. US20020052043A1  
GENERAL INFORMATION:  
APPLICANT: L1, Y1  
APPLICANT: Ruben, Steven M.  
TITLE OF INVENTION: Human G-Protein Coupled Receptors  
FILE REFERENCE: 1488.122003  
CURRENT APPLICATION NUMBER: US/09/827, 937A  
CURRENT FILING DATE: 2001-04-09  
PRIOR APPLICATION NUMBER: 08/852, 824  
PRIOR FILING DATE: 1997-05-07  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 384  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-827-937A-4

Query Match 99.1%; Score 1948; DB 10; Length 384;  
Best Local Similarity 99.0%; Pred. No. 1.7e-157;  
Matches 380; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNATGTPVAPESCQOLAAAGSHRLIVLHYNHSGRLAGRGCPEDGGLGALGSLVAASCLV 60  
1 MNATGTPVAPESCQOLAAAGSHRLIVLHYNHSGRLAGRGCPEDGGLGALGSLVAASCLV 60  
Db 1  
QY 61 VLENLVLAITSHMSRKNVYICLVNTISDLTGAAAYLANVLSARFRLAPQWPL 120  
61 VLENLVLAITSHMSRKNVYICLVNTISDLTGAAAYLANVLSARFRLAPQWPL 120  
Db 1  
QY 121 REGILFTALASTFSLFTAGRFATNVPVPAESGATYTSVYIGICWLLAALGMDP 180  
121 REGILFTALASTFSLFTAGRFATNVPVPAESGATYTSVYIGICWLLAALGMDP 180  
Db 121  
QY 121 RKGILFTALASTFSLFTAGRFATNVPVPAESGATYTSVYIGICWLLAALGMDP 180  
121 RKGILFTALASTFSLFTAGRFATNVPVPAESGATYTSVYIGICWLLAALGMDP 180  
Db 121  
QY 181 LIGMNCACADRCSSLLPLYSKRYILFCLVIFAGVLAITMKGALFRLVQASGQAPRP 240  
181 LIGMNCACADRCSSLLPLYSKRYILFCLVIFAGVLAITMKGALFRLVQASGQAPRP 240  
Db 181  
QY 241 AARRKARRLKTYMLLAFLVCMGPFLGGLLADVFGSNLMAOEYLKGMWILLAVLNS 300  
241 AARRKARRLKTYMLLAFLVCMGPFLGGLLADVFGSNLMAOEYLKGMWILLAVLNS 300  
Db 241  
QY 301 AVNPITISFRSREVCRAVSLFCCGCLRGMGPDCCLARAVEHSGASTTDSLRPRS 360  
301 AVNPITISFRSREVCRAVSLFCCGCLRGMGPDCCLARAVEHSGASTTDSLRPRS 360  
Db 301  
QY 361 FRGSRLSFRMRPLSSSSSVRSI 384  
361 FRGSRLSFRMRPLSSSSSVRSI 384  
Db 361

## RESULT 8

US-09-917-508-2  
Sequence 2, Application US/09917508  
Patent No. US20020137136A1  
GENERAL INFORMATION:  
APPLICANT: TSUI, Ping  
APPLICANT: Ruben, Steven M.  
TITLE OF INVENTION: Rat G Protein Coupled Receptor, EDG6  
FILE REFERENCE: GP-70715  
CURRENT APPLICATION NUMBER: US/09/917, 508  
CURRENT FILING DATE: 2001-07-27  
PRIOR APPLICATION NUMBER: 60/222, 259  
PRIOR FILING DATE: 2000-07-31  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 386

TYPE: PRT  
ORGANISM: RAT  
US-09-917-508-2

Query Match 83.5%; Score 1642; DB 10; Length 386;  
Best Local Similarity 83.1%; Pred. No. 1.2e-131;  
Matches 315; Conservative 29; Mismatches 33; Indels 2; Gaps 2;

QY 6 TPVAPESCQOLAAAGSHRLIVLHYNHSGRLAGRGCPED-GGLGALGSLVAASCLVLEN 64  
6 TPVAPESCQOLAAAGSHRLIVLHYNHSGRLAGRGCPEDNGCLMTGSPVAAGCLVLEN 67  
Db 8  
QY 65 LVLVAITSHMSRKNVYICLVNTISDLTGAAAYLANVLSARFRLAPQWPLRECL 124  
65 LVLVAITSHMSRKNVYICLVNTISDLTGAAAYLANVLSARFRLAPQWPLRECL 124  
Db 68  
QY 125 LFTALASTFSLFTAGRFATNVPVPAESGATYTSVYIGICWLLAALGMDPFLGW 184  
125 LFTALASTFSLFTAGRFATNVPVPAESGATYTSVYIGICWLLAALGMDPFLGW 184  
Db 128  
QY 185 NCLCADRCSSLLPLYSKRYILFCLVIFAGVLAITMKGALFRLVQASGQAPRPARR 244  
185 NCLCADRCSSLLPLYSKRYILFCLVIFAGVLAITMKGALFRLVQASGQAPRPARR 244  
Db 187  
QY 245 KARLLKTYMLLAFLVCMGPFLGGLLADVFGSNLMAOEYLKGMWILLAVLNSAVNP 304  
245 KARLLKTYMLLAFLVCMGPFLGGLLADVFGSNLMAOEYLKGMWILLAVLNSAVNP 304  
Db 247  
QY 305 IISFPRSREVCRAVSLFCCGCLRGMGPDCCLARAVEHSGASTTDSLRPRSFRGS 364  
305 IISFPRSREVCRAVSLFCCGCLRGMGPDCCLARAVEHSGASTTDSLRPRSFRGS 364  
Db 307  
QY 365 RSLSFRMRPLSSSSSVRS 383  
365 RSLSFRMRPLSSSSSVRS 383  
Db 367

## RESULT 9

US-09-731-030A-21  
Sequence 21, Application US/09731030A  
Patent No. US20020142375A1  
GENERAL INFORMATION:  
APPLICANT: MONROE, Donald G.  
APPLICANT: GUPTA, Ashwani K.  
TITLE OF INVENTION: MAMMALIAN EDG-7 RECEPTOR HOMOLOGS  
FILE REFERENCE: 8074-0015  
CURRENT APPLICATION NUMBER: US/09/731, 030A  
CURRENT FILING DATE: 1998-12-29  
PRIOR APPLICATION NUMBER: 60/070, 184  
PRIOR FILING DATE: 1997-12-30  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 21  
LENGTH: 224  
TYPE: PRT  
ORGANISM: Rattus sp.  
US-09-731-030A-21

Query Match 50.5%; Score 993; DB 10; Length 224;  
Best Local Similarity 82.4%; Pred. No. 4.9e-77;  
Matches 182; Conservative 24; Mismatches 15; Indels 0; Gaps 0;

QY 161 RYVGGIGLCMLAALGMDPLGNCCLCAFDRCSSLLPLYSKRYILFCLVIFAGVLAITM 220  
161 RYVGGIGLCMLAALGMDPLGNCCLCAFDRCSSLLPLYSKRYILFCLVIFAGVLAITM 220  
Db 1  
QY 221 GLYGALFRLVQASGQAPRPARRKARLLKTYMLLAFLVCMGPFLGGLLADVFGSNL 280  
221 GLYGALFRLVQASGQAPRPARRKARLLKTYMLLAFLVCMGPFLGGLLADVFGSNL 280  
Db 61  
QY 281 WAOEYLRGMWILLAVLNSAVNPITISFRSREVCRAVSLFCCGCLRGMGPDCCLAR 340  
281 WAOEYLRGMWILLAVLNSAVNPITISFRSREVCRAVSLFCCGCLRGMGPDCCLAR 340

|                       |        |                    |        |            |
|-----------------------|--------|--------------------|--------|------------|
| Query Match           | 32.88; | Score 645.5;       | DB 10; | Length 378 |
| Best Local Similarity | 42.28; | Pred. No. 2.1e-47; |        |            |

Matches 147; Conservative 55; Mismatches 133; Indels 13; Gaps 6;

```
OY 28 HHNSRLAGRGCPEDGGLRGLSVAASCLVYENLVLAITSHMSRRMYIYCLVN 87
    |||:|||||
    21 HYOYVKLAGRLKEASEGSTLTITVLFVICSFTVENLWVLAIRKNNKFHNRMYEFGN 80
OY 88 ITLSDLLTGAAYLANVLGSGARTFLAPAOQFLREGILFTALASTFSLFTAGREFATM 147
    |||:|||||
    81 LALCDLLAGIAVKVNLMSGKFTFSLPTWFLREGSMFVALGASTCSLTAIERHITM 140
OY 148 VRPAESGATKTSRYGFTIGICWLLAALLGMPLGNCNLCAPDGCSSLLPLYSKRTIIF 207
    |||:|||||
    141 IK-MRPYDANKRRHVFLLDGMCMWLAFTLGAALPILGNCNCLHNPDCSTILPLYSKRTIIF 199
OY 208 CLVIFAGVLAITIMGLYGAIFRLVOASQOKAPRPAARRKARLLKTYLMTLLAFVCGMPL 267
    |||:|||||
    200 CISIFPAITLVITVILYARIYFLVSSSRKYANNHNSERSMALLRTVIVIVSVFIACWSPL 259
OY 268 FGLLLADVFSGNLMAOEYLRGMDMILALAVLNSAVNPITISFRSEVCRVAVLSFCGCL 327
    |||:|||||
    260 FILFLIDV-ACRVOACPLIFKAQMFIVLAVLNSAMPVITYTLASEKMRAPFR-LVCNCL 317
OY 328 RLCMRGPGDCLARAVEAHSAGSTDDSLRPRDSFRGSRSLFRMRPEPL 375
    |||:|||||
    318 ---VRGRG---ARASPIQALDPs----RSKSSSSNNSSHSPKXVEDL 355
```

## RESULT 13

```
US-09-731-030A-18
; Sequence 18, Application US/09731030A
; Patent No. US20020142375A1
; GENERAL INFORMATION:
; APPLICANT: MUNROE, Donald G.
; APPLICANT: MUNROE, Donald G.
; APPLICANT: ZASTRANT, Roman L.
; TITLE OF INVENTION: MAMMALIAN EDG-7 RECEPTOR HOMOLOGS
; FILE REFERENCE: 8074-0015
; CURRENT APPLICATION NUMBER: US/09/731,030A
; CURRENT FILING DATE: 1998-12-29
; PRIOR APPLICATION NUMBER: 60/070,184
; PRIOR FILING DATE: 1997-12-30
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Unknown
; OTHER INFORMATION: Description of Unknown Organism: Mammalian EDG
US-09-731-030A-18
```

Query Match 32.8%; Score 645.5; DB 10; Length 378;

Best Local Similarity 42.2%; Pred. No. 2,1e-47;

Matches 147; Conservative 55; Mismatches 133; Indels 13; Gaps 6;

```
OY 28 HHNSRLAGRGCPEDGGLRGLSVAASCLVYENLVLAITSHMSRRMYIYCLVN 87
    |||:|||||
    21 HYOYVKLAGRLKEASEGSTLTITVLFVICSFTVENLWVLAIRKNNKFHNRMYEFGN 80
OY 88 ITLSDLLTGAAYLANVLGSGARTFLAPAOQFLREGILFTALASTFSLFTAGREFATM 147
    |||:|||||
    81 LALCDLLAGIAVKVNLMSGKFTFSLPTWFLREGSMFVALGASTCSLTAIERHITM 140
OY 148 VRPAESGATKTSRYGFTIGICWLLAALLGMPLGNCNLCAPDGCSSLLPLYSKRTIIF 207
    |||:|||||
    141 IK-MRPYDANKRRHVFLLDGMCMWLAFTLGAALPILGNCNCLHNPDCSTILPLYSKRTIIF 199
OY 208 CLVIFAGVLAITIMGLYGAIFRLVOASQOKAPRPAARRKARLLKTYLMTLLAFVCGMPL 267
    |||:|||||
    200 CISIFPAITLVITVILYARIYFLVSSSRKYANNHNSERSMALLRTVIVIVSVFIACWSPL 259
OY 268 FGLLLADVFSGNLMAOEYLRGMDMILALAVLNSAVNPITISFRSEVCRVAVLSFCGCL 327
    |||:|||||
    260 FILFLIDV-ACRVOACPLIFKAQMFIVLAVLNSAMPVITYTLASEKMRAPFR-LVCNCL 317
```

DB 260 FILFLIDV-ACRVOACPLIFKAQMFIVLAVLNSAMPVITYTLASEKMRAPFR-LVCNCL 317

OY 328 RLCMRGPGDCLARAVEAHSAGSTDDSLRPRDSFRGSRSLFRMRPEPL 375

DB 318 ---VRGRG---ARASPIQALDPs----RSKSSSSNNSSHSPKXVEDL 355

## RESULT 14

```
US-10-037-616-21
; Sequence 21, Application US/10037616
; Patent No. US2002013148A1
; GENERAL INFORMATION:
; APPLICANT: English, Denis
; APPLICANT: Kovacs, Richard J.
; APPLICANT: Rizzo, Maria T.
; APPLICANT: Silva, Daniel T.
; TITLE OF INVENTION: Sphingolipid compositions and methods for their therapeutic use
; FILE REFERENCE: 7042-119
; CURRENT APPLICATION NUMBER: US/10/037,616
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/243,887
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-037-616-21
```

Query Match 32.8%; Score 645.5; DB 12; Length 378;

Best Local Similarity 42.2%; Pred. No. 2,1e-47;

Matches 147; Conservative 55; Mismatches 133; Indels 13; Gaps 6;

```
OY 28 HHNSRLAGRGCPEDGGLRGLSVAASCLVYENLVLAITSHMSRRMYIYCLVN 87
    |||:|||||
    21 HYOYVKLAGRLKEASEGSTLTITVLFVICSFTVENLWVLAIRKNNKFHNRMYEFGN 80
OY 88 ITLSDLLTGAAYLANVLGSGARTFLAPAOQFLREGILFTALASTFSLFTAGREFATM 147
    |||:|||||
    81 LALCDLLAGIAVKVNLMSGKFTFSLPTWFLREGSMFVALGASTCSLTAIERHITM 140
OY 148 VRPAESGATKTSRYGFTIGICWLLAALLGMPLGNCNLCAPDGCSSLLPLYSKRTIIF 207
    |||:|||||
    141 IK-MRPYDANKRRHVFLLDGMCMWLAFTLGAALPILGNCNCLHNPDCSTILPLYSKRTIIF 199
OY 208 CLVIFAGVLAITIMGLYGAIFRLVOASQOKAPRPAARRKARLLKTYLMTLLAFVCGMPL 267
    |||:|||||
    200 CISIFPAITLVITVILYARIYFLVSSSRKYANNHNSERSMALLRTVIVIVSVFIACWSPL 259
OY 268 FGLLLADVFSGNLMAOEYLRGMDMILALAVLNSAVNPITISFRSEVCRVAVLSFCGCL 327
    |||:|||||
    260 FILFLIDV-ACRVOACPLIFKAQMFIVLAVLNSAMPVITYTLASEKMRAPFR-LVCNCL 317
OY 328 RLCMRGPGDCLARAVEAHSAGSTDDSLRPRDSFRGSRSLFRMRPEPL 375
    |||:|||||
    318 ---VRGRG---ARASPIQALDPs----RSKSSSSNNSSHSPKXVEDL 355
```

## RESULT 15

```
US-09-815-333-2
; Sequence 2, Application US/09815333
; Patent No. US20010041355A1
; GENERAL INFORMATION:
; APPLICANT: Ramakrishnan, Shyam
; TITLE OF INVENTION: Regulation of Human Nerve Growth
; FILE REFERENCE: 04974,00461
; CURRENT APPLICATION NUMBER: US/09/815,333
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/191,766
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 5
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-815-333-2

```

```

Query Match          32.8%; Score 645.5; DB 10; Length 398;
Best Local Similarity 40.8%; Pred. No. 2,2e-47;
Matches 154; Conservative 50; Mismatches 126; Indels 47; Gaps 7;

```

```

QY 22 SRLIVLHYHNSGRLAGRGPEDGGIGALRGLSVAASCLVLENLVLAITSHMRSRRWV 81
   | : ||||| : || : : ||||| || : | |
Db 12 SEVIYVLYHNYTGRLRGARYQPGAGLRADAVCIACAFIVLENLAVLVIGRHPFHAPM 71
   : | : ||||| ||||| ||||| : ||||| ||||| : ||||| |||||
QY 82 YYCLVITLSDLTGAAVLANVLLSGARFERRIAPQMFIREGLFTALASTFSLFTAG 141
   : | : ||||| ||||| ||||| : ||||| ||||| : ||||| |||||
Db 72 FLLLSGLTSLDLAGAAVLANVLLSGPLTLKLSPALMFAREGGVFVALTASVLSLAIATL 131
   : | : ||||| ||||| ||||| : ||||| ||||| : ||||| |||||
QY 142 ERFATWVR---PVASGATKTSRVYGFILGWLALILGMLPLIGWNCLCAFDRCSLL 197
   || || | || : | | | | : ||||| ||||| ||||| |||||
Db 132 ERSLTWARRGPAVSSRG-----RTLMAAAMGVSLILGLPALGWNCLGRIDACSTVL 186
   || || | || : | | | | : ||||| ||||| ||||| |||||
QY 198 PLYSKRYIIFCLVIFAGVLTATMGLYGATFRLVQASGQKAP-RPA-----ARRKAR- 247
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db 187 PLYAKAYVLEFCVLAFAVGILAAICALYARIYCOVRANARRLPARPGTGTSTRARRKPRS 246
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
QY 248 -RLTKTVLMLLAFVLCWGPFLGLLADYFGSNIMAOEYLRGMWILALAVLNSAVNPFI 306
   ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 LALIKTLISVLLAFVACWGPFLFLLLLDVACPARTCPVILQA-DPFLGLAMANSLLNPFI 305
   ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 YSFRSREVCRAVLSFLCG-----CLRLGMRGPGDCLAR 340
   | : || : | : |||||
Db 306 YTLTNRDLRHALLRLVCCGRHSCGRDPSSGQGSASAAEASGGLARRCLPPGLDGSFSSSER 365
   || : || : | : |||||
QY 341 AVEAHSGASTDSSLRP 357
   : | : || : | : |||||
Db 366 SSPQRDGLDTSGSTGSP 382
   : | : || : | : |||||

```

Search completed: December 9, 2002, 12:06:00  
 Job time : 13 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 12:03:06 ; Search time 16 Seconds  
(without alignments)  
706.150 Million cell updates/sec

Title: US-09-786-926-1

Perfect score: 1966  
Sequence: 1 MNATGTPVAPESCCQLAAG.....RSLSPMRPLSSISVRSI 384

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCtus.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID               | Description       |
|------------|-------|-------------|--------|---------------------|-------------------|
| 1          | 1948  | 99.1        | 384    | 3 US-08-852-824-4   | Sequence 4, Appl  |
| 2          | 1437  | 73.1        | 509    | 2 US-08-845-566-1   | Sequence 1, Appl  |
| 3          | 666   | 33.9        | 383    | 1 US-08-196-989B-4  | Sequence 4, Appl  |
| 4          | 666   | 33.9        | 383    | 2 US-08-760-936-4   | Sequence 2, Appl  |
| 5          | 653   | 33.2        | 382    | 4 US-09-542-733-2   | Sequence 2, Appl  |
| 6          | 645.5 | 32.8        | 378    | 4 US-09-082-088-2   | Sequence 2, Appl  |
| 7          | 645.5 | 32.8        | 378    | 4 US-09-546-117-2   | Sequence 2, Appl  |
| 8          | 635.5 | 32.3        | 382    | 4 US-09-262-477-2   | Sequence 2, Appl  |
| 9          | 633   | 32.2        | 381    | 2 US-08-845-566-3   | Sequence 2, Appl  |
| 10         | 633   | 32.2        | 381    | 2 US-08-467-948A-28 | Sequence 3, Appl  |
| 11         | 633   | 32.2        | 381    | 3 US-08-852-824-18  | Sequence 28, Appl |
| 12         | 633   | 32.2        | 381    | 3 US-08-467-947A-28 | Sequence 28, Appl |
| 13         | 633   | 32.2        | 381    | 5 PCT-US96-10618-4  | Sequence 4, Appl  |
| 14         | 600   | 30.5        | 334    | 1 US-08-118-270-73  | Sequence 73, Appl |
| 15         | 600   | 30.5        | 334    | 5 PCT-US93-08528-73 | Sequence 73, Appl |
| 16         | 517.5 | 26.3        | 352    | 1 US-08-196-989B-2  | Sequence 2, Appl  |
| 17         | 517.5 | 26.3        | 352    | 1 US-08-760-936-2   | Sequence 2, Appl  |
| 18         | 479.5 | 24.4        | 351    | 3 US-08-789-982-2   | Sequence 2, Appl  |
| 19         | 476.5 | 24.2        | 351    | 3 US-08-861-747-2   | Sequence 2, Appl  |
| 20         | 444.5 | 22.6        | 364    | 4 US-08-763-938-2   | Sequence 2, Appl  |
| 21         | 439   | 22.3        | 393    | 5 PCT-US96-10618-3  | Sequence 3, Appl  |
| 22         | 434   | 22.1        | 364    | 5 PCT-US96-10618-2  | Sequence 2, Appl  |
| 23         | 418   | 21.3        | 354    | 4 US-09-325-997-2   | Sequence 2, Appl  |
| 24         | 418   | 21.0        | 354    | 4 US-09-837-726-2   | Sequence 2, Appl  |
| 25         | 413.5 | 21.0        | 353    | 3 US-08-997-803-14  | Sequence 14, Appl |
| 26         | 334.5 | 17.0        | 213    | 3 US-08-997-803-15  | Sequence 15, Appl |
| 27         | 298   | 15.2        | 345    | 1 US-08-118-270-70  | Sequence 70, Appl |

|    |       |      |     |                     |                   |
|----|-------|------|-----|---------------------|-------------------|
| 28 | 298   | 15.2 | 345 | 5 PCT-US93-08528-70 | Sequence 70, Appl |
| 29 | 295.5 | 15.0 | 393 | 2 US-08-467-948A-4  | Sequence 4, Appl  |
| 30 | 295.5 | 15.0 | 393 | 3 US-08-467-947A-4  | Sequence 4, Appl  |
| 31 | 254   | 12.9 | 775 | 4 US-09-513-838-6   | Sequence 6, Appl  |
| 32 | 248.5 | 12.6 | 362 | 4 US-09-513-838-2   | Sequence 2, Appl  |
| 33 | 248.5 | 12.4 | 412 | 1 US-08-349-696-21  | Sequence 21, Appl |
| 34 | 244.5 | 12.4 | 412 | 1 US-08-233-009-21  | Sequence 21, Appl |
| 35 | 244.5 | 12.4 | 412 | 2 US-08-560-231-21  | Sequence 21, Appl |
| 36 | 244.5 | 12.4 | 412 | 4 US-09-080-704A-21 | Sequence 21, Appl |
| 37 | 240.5 | 12.2 | 325 | 4 US-08-387-805-16  | Sequence 16, Appl |
| 38 | 235   | 12.0 | 407 | 1 US-08-293-563-6   | Sequence 6, Appl  |
| 39 | 228   | 11.6 | 572 | 1 US-08-334-698-2   | Sequence 2, Appl  |
| 40 | 228   | 11.6 | 572 | 1 US-08-228-932-2   | Sequence 2, Appl  |
| 41 | 228   | 11.6 | 572 | 1 US-08-468-939-2   | Sequence 2, Appl  |
| 42 | 228   | 11.6 | 572 | 1 US-08-722-001-30  | Sequence 30, Appl |
| 43 | 228   | 11.6 | 572 | 1 US-08-406-855A-2  | Sequence 2, Appl  |
| 44 | 228   | 11.6 | 572 | 2 US-08-722-190-2   | Sequence 2, Appl  |
| 45 | 228   | 11.6 | 572 | 3 US-08-244-354-2   | Sequence 2, Appl  |

ALIGNMENTS

RESULT 1  
US-08-852-824-4  
; Sequence 4, Application US/08852824C  
; Patent No. 6060272  
; GENERAL INFORMATION:  
; APPLICANT: li et al.  
; TITLE OF INVENTION: Human G-Protein Coupled Receptors  
; FILE REFERENCE: 1488, 1220000  
; CURRENT APPLICATION NUMBER: US/08/852, 824C  
; CURRENT FILING DATE: 1997-05-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: genomic  
US-08-852-824-4

|                       |                 |   |           |             |
|-----------------------|-----------------|---|-----------|-------------|
| Query Match           | 99.1%;          | Score 1948;   | DB 3;     | Length 384; |
| Best Local Similarity | 99.0%;          | Pred. No. 4.1e-172;                                       |           |             |
| Matches 380;          | Conservative 3; | Mismatches 1;   | Indels 0; | Gaps 0;     |
| QY                    | 1               | MNATGTPVAPESCCQLAAGSHRLVLYHNSGRLAGRGPEDEGLAGLSVAASCTV     | 60        |             |
| DB                    | 1               | MNATGTPVAPESCCQLAAGSHRLVLYHNSGRLAGRGPEDEGLAGLSVAASCTV     | 60        |             |
| QY                    | 61              | VLENILVLAITSHMRRRVYVCIVNTISDLTGAAIYANVLSGARTFLAPQWFL      | 120       |             |
| DB                    | 61              | VLENILVLAITSHMRRRVYVCIVNTISDLTGAAIYANVLSGARTFLAPQWFL      | 120       |             |
| QY                    | 121             | REGILFLTAALSTFSLFTAGERFATWRPVAESGATKTSRYVGFGLCMLAALIGMPL  | 180       |             |
| DB                    | 121             | RKGLFLTAALSTFSLFTAGERFATWRPVAESGATKTSRYVGFGLCMLAALIGMPL   | 180       |             |
| QY                    | 181             | LLGNMNCACDRSSLLPLYSKRYILFCVITFAGVATIMGLYGAIFRLVQASGQKAPRP | 240       |             |
| DB                    | 181             | LLGNMNCACDRSSLLPLYSKRYILFCVITFAGVATIMGLYGAIFRLVQASGQKAPRP | 240       |             |
| QY                    | 241             | AARRKARRLKTYMLITLAFVWGPFLGILLADVGSNLMQEYLRGMDITLAVLNS     | 300       |             |
| DB                    | 241             | AARRKARRLKTYMLITLAFVWGPFLGILLADVGSNLMQEYLRGMDITLAVLNS     | 300       |             |
| QY                    | 301             | AVNPITYSFRSREVCRAVLSFLCCGLRLGMRGPGDCLRAVEAHSASTTDSLRPRS   | 360       |             |
| DB                    | 301             | AVNPITYSFRSREVCRAVLSFLCCGLRLGMRGPGDCLRAVEAHSASTTDSLRPRS   | 360       |             |
| QY                    | 361             | FRGSRSLSPMRPLSSISVRSI 384                                 |           |             |
| DB                    | 361             | FRGSRSLSPMRPLSSISVRSI 384                                 |           |             |

## RESULT 2

US-08-845-566-1

; Sequence 1, Application US/08845566

; Patent No. 5912144

; GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice

; APPLICANT: Guegler, Karl

; TITLE OF INVENTION: EDG-1 LIKE RECEPTOR

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: IBM Compatible

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/845,566

; FILING DATE: Filed Herewith

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0271 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 509 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: TYMNOR01

; CLONE: 144690

; US-08-845-566-1

Query Match 73.1%; Score 1437; DB 2; Length 509;

Best Local Similarity 76.3%; Pred. No. 8,6e-125;

Matches 303; Conservative 8; Mismatches 42; Indels 44; Gaps 6;

QY 1 MNATGPVAPESCOOLAAGHSRLVLYLHNHSGRLAGRGPEDEGLGALRGSLVSAASCLV 60  
DB 1 MNATGPVAPESCOOLAAGHSRLVLYLHNHSGRLAGRGPEDEGLGALRGSLVSAASCLV 60  
QY 61 VLENLIVLAATSHMRSRWVYVCLVNTLSDLLTGAAYLANVLLSGARTFRLAPAQWFL 120  
DB 61 VLENLIVLAATSHMRSRWVYVCLVNTLSDLLTGAAYLANVLLSGARTFRLAPAQWFL 120  
QY 121 REGILFTALAASTFSLFTAGEFATWVRPVASGATKTSRYVGFGLCWLALAILGMLP 180  
DB 121 REGILFTALAASTFSLFTAGEFATWVRPVASGATKTSRYVGFGLCWLALAILGMLP 180  
QY 181 LIGMNCICAFDRCSLLPLYSKRYILFCVITFAGVATIMGLYGAIFRLVQASGOKAPRP 240  
DB 181 LIGMNCICAFDRCSLLPLYSKRYILFCVITFAGVATIMGLYGAIFRLVQASGOKAPRP 240  
QY 241 AARRKARRLKTVLMTLLAFVCMGPELFGILLADVFCSNLMAG----- 283  
DB 241 AARRKARRLKTVLMTLLAFVCMGPELFGILLADVFCSNLMAG----- 283  
QY 284 -EYLKGMDFLALAVLNSAVNPITYSFRSREVCRAVLSPGCCGLRGLMGRGPDCLARAY 342  
DB 284 -EYLKGMDFLALAVLNSAVNPITYSFRSREVCRAVLSPGCCGLRGLMGRGPDCLARAY 342  
QY 297 RPQLGGPHTLLPQGGVGS-----RAQPSAAGVSGWAC-----EGPDCLARAY 343  
DB 297 RPQLGGPHTLLPQGGVGS-----RAQPSAAGVSGWAC-----EGPDCLARAY 343

QY 343 EAHSGASTDSSLRPRDSRSLSFRMRPLSSIS 379

DB 344 EAHSGA-----SNHRQLSEAKGQLS-----RLPLAQLS 371

## RESULT 3

US-08-196-989B-4

; Sequence 4, Application US/08196989B

; Patent No. 5585476

; GENERAL INFORMATION:

; APPLICANT: MacLennan, A. John

; TITLE OF INVENTION: Molecular Cloning and Expression of

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Saliwanchik &amp; Saliwanchik

; STREET: 2421 N.W. 41st Street, Suite A-1

; CITY: Gainesville

; STATE: FL

; COUNTRY: US

; ZIP: 32606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/196,989B

; FILING DATE: 15-FEB-1994

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Lloyd, Jeff

; REGISTRATION NUMBER: 35,589

; REFERENCE/DOCKET NUMBER: MAC-100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 904-375-8100

; TELEFAX: 904-372-5800

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 383 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-196-989B-4

Query Match 33.9%; Score 666; DB 1; Length 383;

Best Local Similarity 40.0%; Pred. No. 9.4e-54;

Matches 150; Conservative 69; Mismatches 126; Indels 30; Gaps 8;

QY 1 MNATGPVAPESCOOLAAGHSRLVLYLHNHSGRLAGRGPEDEGLGALRGSLVSAASCLV 60  
DB 2 VSSSTIPVAKLRSGVSDYGVNVDIVRHNYTGKL-NIGVEKDHGILTSVFTLICLLI 60  
QY 61 VLENLIVLAATSHMRSRWVYVCLVNTLSDLLTGAAYLANVLLSGARTFRLAPAQWFL 120  
DB 61 ILENLIVLLTITWKTKFRPMYVYFIGNLALSDLLAGVAYTNLLLSGATYTKLTPAQWFL 120  
QY 121 REGILFTALAASTFSLFTAGEFATWVRPVASGATKTSRYVGFGLCWLALAILGMLP 180  
DB 121 REGSFVVALSVSLSLAIAERITIMKMLHNG-SNSSSFLLISACWVLSIILGGLP 179  
QY 181 LIGMNCICAFDRCSLLPLYSKRYILFCVITFAGVATIMGLYGAIFRLVQASGOKAPRP 240  
DB 180 IMGNMNCISLSSGCVTLPLVYKHVYILFCVITFGLLLSIYLVLCRIYSLVTRSRULFR 239  
QY 236 ----KAPRPARRKARRLKTVLMTLLAFVCMGPELFGILLADVFCSNLMAG----- 291  
DB 240 KNISRAAR-SSEKSLALKTVIIVLSYFIACWAPLFLILLDDV-GCKATCTDILYAEY 296  
QY 292 IIALAVLNSAVNPITYSFRSREVCRAVLSPGCCGLRGLMGRGPDCLARAY 343  
DB 297 FLVLAVLNSGPNPIITYITLNKEMRAFIIRITISCKCPNGDSAGNFKRPILPGMEFSR--- 353

```
QY      344 AHGCASTDSSLRPR 358
          | : | : | :
Db      354 ----SKSDNSSHPQ 363
```

#### RESULT 4

Sequence 4, Application US/08760936  
Patent No. 5856443  
GENERAL INFORMATION:  
APPLICANT: MacLennan, A. John  
TITLE OF INVENTION: Molecular Cloning and Expression of  
TITLE OF INVENTION: G-Protein Coupled Receptors  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Salliwanchik, Lloyd & Salliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: US  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/760,936  
FILING DATE: December 6, 1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Doran R.  
REGISTRATION NUMBER: 38,261  
REFERENCE/DOCKET NUMBER: MAC-100C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-372-5800  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 383 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-760-936-4

|             |        |            |       |             |
|-------------|--------|------------|-------|-------------|
| Query Match | 33.98; | Score 666; | DB 2; | Length 383; |
|-------------|--------|------------|-------|-------------|

```

QY 1 MNATGFVAPEESOCQIAAGGSRLLVLYHNHSGRIAGGGEDEGIALRELTSVAACIV 60
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 VSSSTIPYVKALASQYSDYDGNDIIVRHNTYGRK - NIGVEKDHGINKLTSVFLICCL 60
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 VLENRLVLAATISHMSRRWYTCVLNITLSDLTGAAIYLANVLLSGARFFRLAPAOEVL 120
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 ILEHIFLLITWTKKFRHPMYEFIGNALSDLLAGVATYNTLLISGTYTKYLPAPMFL 120
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 REGLEFALAAAYFSLTFAGERFATWAPVAESGATSRVYGFIGLCWMLAALLGMPL 180
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 REGSMFALTSAYFSLTALAIERYITMLKMKLHNG - SNSSSFLLISACWVLSILIGLP 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 LLGMCILCAFDRCSSLLPLYSKRYIFCLVIFAGVLAITIMGLYCAIFRLYQASQ----- 235
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 IMGNCSISLSSGSLVPLYHKRYITFCTYVTLTLLISIVLYCRYISLVNTRSRLLTFR 239
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 236 ----KAPRPARRKARVRLKTVMLLITLAVLYCWGDFELGLLADVGSMLMAOEYLRGMDV 291
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 KNISKASR--SSEKSLATLKTIVLYLVSPLTACMAPLFLITLLDV -GCKAKTCOILYKAEY 296
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 292 ILALAVYANSVNPITYSPRSREYCAVSEFL -CCGCLRLGMRG-----PGDCIQAAYE 343
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 FLVYIAVANSQNPITYTLNKKERRRFEIRIISCCCKCPGDSAGKFRRIIPGMEFSR-- 353
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```
QY      344 AHSGASTTDDSSLRPR 358
          | : | : | :
Db      354 ----SKSDNSSHPQ 363
```

## RESULT

```

1 Sequence 2, Application US/095542733
2 Patient No. 6323333
3
4 GENERAL INFORMATION:
5
6 APPLICANT: Ping Tsui
7
8 TITLE OF INVENTION: Mouse EdG1
9
10 FILE REFERENCE: GP-70610
11
12 CURRENT APPLICATION NUMBER: US/09/542,733-04-04
13
14 CURRENT FILING DATE: 2000-04-04
15
16 EARLIER APPLICATION NUMBER: 60/127,696
17
18 EARLIER FILING DATE: 1999-04-05
19
20 NUMBER OF SEQ ID NOS: 2
21
22 SOFTWARE: FastSeq for Windows Version 3.0
23
24 SEQ ID NO 2
25
26 LENGTH: 382
27
28 TYPE: PRT
29
30 ORGANISM: MUS MUSCULUS
31
32 US-09-542-733-2

```

|                       |        |            |        |             |
|-----------------------|--------|------------|--------|-------------|
| Query Match           | 33.28; | Score 653; | DB 4;  | Length 382; |
| Post Local similarity | 30.59; | Prod NO. 1 | 50-53; |             |

```

QY      1  MNANGTVEAEEGCOOLAAAGHSLLVLAHVNNHSGRLAGGCGEDGITALAGLSVAASCLV 60
Db      1  MVTSEIPEVALNSSVDYGNVDIIVRHNYNGKL-NIGAEKDHGDKLTSVPELLICCF 59
QY      61  VLEMLVLAITSBMSRRMYVLCVNTVITLSLTGAVLVANVLLSGAFETRLPAQWFL 120
Db      60  ILEMTIFELTWTGKTKFFHRPMYTFIGNLALSPLLAGVAYTMNLLSGATTYKRLTPAQWL 119
QY      121  REGLEFALAASTFSLFTAGERFATMPVPAESGATYTSRYVCFIGLCWLLALLGMDP 180
Db      120  REGMFPAALASVFSLLAIAIERIYTMKLMKLRHNG-SNSSSFLLISACWVTSLLIGCP 178
QY      181  ILGNCCACDPROSSLPIFYSKRIYLCFVIAFVAGVATIMLYGAFIRIVQASGO----- 235
Db      179  IMGNCSSISSGCVYPLPHKHXYLLFCQYVFTLLLSIVLYLCRIYISIVRTSRRLTFR 238
QY      236  ---KAPRPARRKARRLKTYIMLILAFVLCVWGCLFGLLADVPGSILNAOETILROMDY 251
Db      235  KNISKASR--SSEKSLALLTIVITVIVSFIACWAPLFTLLILDV-GCKAKTCDITLKAEY 255
QY      292  ILALAVINSAVNPIITYSFRSEYCAVLAFL--CCGCLRLGNHG-----PGCDLARAVE 343
Db      296  FLVLAIVNNGNPILYITLTNKEKRRARFIYVSCGCKCPNGDSAGKFKRPIIPGMEFSR--- 352
QY      344  AHSCASTTDSLSLPR 358
Db      353  -----SKSDNSGHPQ 362

```

## RESULT

US-09-082-088-2  
Sequence 2, Application US/09082088  
Patent No. 6130067  
GENERAL INFORMATION:  
APPLICANT: TSUI, PING  
TITLE OF INVENTION: HUMAN EDG3sb GENES  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rater & Prestia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA

```

ZIP : 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,088
FILING DATE: 20-MAY-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GP-70453
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX:
INFORMATION FOR SEQ ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-082-088-2

Query Match      32.8%; Score 645.5; DB 4; Length 378;
Best local Similarity 42.2%; Pred. No. 7.2e-52;
Matches 147; Conservative 55; Mismatches 133; Indels 13; Gaps 6;

QY   28 HYNHSGRLAGRGGEDEGGALGRLGSVAASCLVIVENLTVLAATISMRSRRWYYCLVN 87
    || :||||| | | | | :|||:||| | | | | | | | | | | | | | | | |
Db   21 HYQVGKGLAGRLKEAEGESTITFTVFYICSPYIENIMVLAIWKRNKKHNMYFIQN 80
    : ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   88 ITSDLLTGAAITYLVNLSGARTFRLPAPOMFLREGLLFTALAISTSLFTYGERPATM 147
    : ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   81 LALCDILAGIAYVNVNLMSGRTFSLSPTWFPLRGSMTFVALGSTSLSLAIAIERHTLM 140
    : ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   148 VRPAESGATKTSTRVYGFICLCMLIALALIGMLPLDGMNCCLCAFDPRCSLLPDKSKRYTF 207
    :: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   141 IK-WRPYDANKRRARVFLILGMCMILAIFLAGLPIFGNMCLHLNDCTILPLYSKIYINF 199
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   208 CLVIYFACVLTIMGLYGAFIRLVQASGOKAPRAPARRARKARLLTVMIIALFYCWGPRL 267
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   200 CISTFAILTVTYIILARIYFLVKSSSRKVANNHNNSESMALLTTVVIVSVETIACWSPL 259
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   268 FGLILLADVEGNSLMAOEYLRGMDYLALAVLNASAVNPITYSFGRSECRALVFFLICCGCL 327
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   260 FILEFLIDV--ACRVAQCPLIFKAQWETIVLAVLNAMNPITYVLASKEMRRAPFR-IYONCL 317
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   328 RLGMRGPGDCILAAVFAHSHGASTYTSSLLPRDPDFSGRSLSFRRKREPL 375
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   318 ---VRGRG--ARASPTQPALDPS---RSKSSSSSNSSHSHPKVEDL 355
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-09-546-117-2
Sequence 2, Application US/09546117
Patent No. 6344542
GENERAL INFORMATION:
APPLICANT: TSUI, PTNG
TITLE OF INVENTION: HUMAN EDG3d GENE
FILE REFERENCE: GP-70453-D1
CURRENT APPLICATION NUMBER: US/09/546,117
CURRENT FILING DATE: 2000-04-10
PRIOR APPLICATION NUMBER: 09/082,088
PRIOR FILING DATE: 1998-05-20
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FASTSEQ for Windows Version 3.0
```

```

; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRY
; ORGANISM: HOMO SAPIENS
US-09-546-117-2

Query Match                      32.8%; Score 645.5; DB 4; Length 378;
Best Local Similarity 42.2%; Pred. No. 7,2e-52;
Matches 147; Conservative 55; Mismatches 133; Indels 13; Gaps 6;

OY      28  HYNSSGLAAGCGEDGCLALRGSLVAASCLVLENNILVAATNSHRSRBRWYYCIVN 87
        || : |::|||: | : | : |::|||: | : | : | : | : | : | : | : |
Db       21  HYQYKLAGRLKRASGSSTLTTLVLVCSTFVENLWLAIKKNKFPINRKYFTIGN 80
        : |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
OY      88  TTLSDELTGAAVLANVLLSGARTFLPAPQWPLREGLFTALASTFSLFTAGEPATM 147
        : |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
Db       81  LAICDLLAGIAVKVNIIIMSGSKTFESLPYVWFREGSMFEVALGASTCSLIAIAERHLT 140
        : |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
OY      148  VRPAESGATKTSRVYGFTGCMLLAALLGMPLPGMNCLCAFDSCSLPLYSKRYTLF 207
        : |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
Db       141  IK-WRPDPANKRRHRVPILLICMCWLIAPFTIGALPILGNCHNLHPDCSTITLPLYSKKITAF 199
        : |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
OY      208  CLVIPAGVLTIMGTCATFRIVQAASGOKAPRAPARRARLKLTVMILTILAFLVCMP 267
        C I:SIFETIALVTLYLVLRITFYELVKSSSRVYAHNNSESMALLRTVYLVSVFACMSPL 259
        : |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
OY      268  FGLIADVFGSNLMAQCYLGRMDMITALAVINSANVPITYSPRSREVCRAVLFCCGCL 327
        |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
Db       260  FILPLIDY-ACRVQACTILFKAQWFTIVLANVSNMNPVITYTLASSEMRAFPR-LYCNC 317
        : |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
OY      328  RLGMRGPDCLARAWEASHGASTTDSSLRPDRDSFRGNSLSFRMRPEPL 375
        : |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
Db       318  --VRGRG--ARASPTQPALDPS---RSKSSSSNSSSHSPKREDL 355

RESULT 8
US-09-262-477-2
; Sequence 2, Application US/09262477A
; Patent No. 6423508
; GENERAL INFORMATION:
; APPLICANT: George Livi
; APPLICANT: Derk Bergama
; APPLICANT: Jeffrey Stadel
; APPLICANT: Winnie Chan
; APPLICANT: Randall Johnson
; APPLICANT: Shelagh Wilson
; APPLICANT: Jon Chambers
; APPLICANT: Philippe Robert
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDE SEQUENCES OF HUMAN EDG-1C AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR AGONISTS AND ANTAGONISTS FOR HUMAN
; TITLE OF INVENTION: EDG-1C RECEPTOR AND SPHINGOSINE 1-PHOSPHATE LIGAND AND
; TITLE OF INVENTION: INTERACTION THEREOF
; FILE REFERENCE: GP50006
; CURRENT APPLICATION NUMBER: US/09/262,477A
; CURRENT FILING DATE: 1999-03-04
; EARLIER APPLICATION NUMBER: 60/077,369
; EARLIER FILING DATE: 1998-03-09
; EARLIER APPLICATION NUMBER: 60/087,102
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 382
; TYPE: PRY
; ORGANISM: Human
US-09-262-477-2

Query Match                      32.3%; Score 635.5; DB 4; Length 382;
Best Local Similarity 40.5%; Pred. No. 6,1e-51;
Matches 136; Conservative 64; Mismatches 121; Indels 15; Gaps 6;

1 MNATGTVPAPESCOQLAAGHSRLIVLHNHSGRLAAGCGEPDDGCLALRGSLVAASCLV 60

```

Db 1 MGPTSVPLVKAHRSVSDVYVNDIIVRHNYTGKL-NISADKENSIKLTSVFFILICCFI 59  
QY 61 VLENLLVLAITSHMRSRRMYCCLVNTITSDLTGAAYLANVLISGARTRLAPAQWFL 120  
Db 60 ILENIFVLLTIWTKRKHRRPMYFIGNLALSDLAGAVATYANLLISGATTYKLPQWFL 119  
QY 121 REGILFTALASTFSLFTAGERPAWVPAVSGATKTSRYGFGICWLLAALIGMLP 180  
Db 120 REGSMFVALSASFSLAIATERYITMLKMKLHNG-SNNFRFLILISACWVLSILIGSLP 178  
QY 181 LIGMNCIAFDRCSSLLPLYSKRYLLFCVIFAGVATIMGLYGAIFRLVOASGQ----- 235  
Db 179 IMGWNCISALSSCSTVPLHKKHYLLFCTVTFVTLTLLSVILYCRISLVTRRSRLTFR 238  
QY 236 ----AAPPRAARKARKRLIKTYMLLAFVCMGPFLLADVFGSNMAOEYLRGMW 291  
Db 239 KNISKASR--SEKSLALKTVIIVLSVFIACWAPLFIILLDV-GCKVKTCIDILFRAEY 295  
QY 292 ILALAVLSAVNPITYSFRSEVCAVLSFL-CCGC 326  
Db 296 FLVLAVLSGNPIIYTLTKEMRAAFIRIMSCCKC 331

## RESULT 9

US-08-845-566-3  
; Sequence 3, Application US/08845566  
; Patent No. 5912144  
; GENERAL INFORMATION:  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: EDG-1 LIKE RECEPTOR  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174, Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/845,566  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0271 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 381 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: Genbank  
; CLONE: 181948  
; US-08-845-566-3

Query Match 32.2%; Score 633; DB 2; Length 381;  
Best Local Similarity 40.5%; Pred. No 1e-50;  
Matches 135; Conservative 63; Mismatches 125; Indels 10; Gaps 5;

QY 1 MNATGPVAPESQOQLAAGHSRLIVLHNHSGRLAGRGSPEDGIGALRGLSVAASCLV 60  
Db 1 MGPTSVPLVKAHRSVSDVYVNDIIVRHNYTGKL-NISADKENSIKLTSVFFILICCFI 59  
QY 61 VLENLLVLAITSHMRSRRMYCCLVNTITSDLTGAAYLANVLISGARTRLAPAQWFL 120  
Db 60 ILENIFVLLTIWTKRKHRRPMYFIGNLALSDLAGAVATYANLLISGATTYKLPQWFL 119  
QY 121 REGILFTALASTFSLFTAGERPAWVPAVSGATKTSRYGFGICWLLAALIGMLP 180  
Db 120 REGSMFVALSASFSLAIATERYITMLKMKLHNG-SNNFRFLILISACWVLSILIGSLP 178  
QY 181 LIGMNCIAFDRCSSLLPLYSKRYLLFCVIFAGVATIMGLYGAIFRLVOASGQAKAPP 240  
Db 179 IMGWNCISALSSCSTVPLHKKHYLLFCTVTFVTLTLLSVILYCRISLVTRRSRLTFR 238  
QY 241 AARRKAR-----LAKTVIMILAFVCMGPFLLADVFGSNMAOEYLRGMW 294  
Db 239 KNISKASRSEBVALKTVIIVLSVFIACWAPLFIILLDV-GCKVKTCIDILFRAEYFLV 297  
QY 295 LAVYLSAVNPITYSFRSEVCAVLSFL-CCGC 326  
Db 298 LAVLSGNPIIYTLTKEMRAAFIRIMSCCKC 330

## RESULT 10

US-08-467-948A-28  
; Sequence 28, Application US/08467948A  
; Patent No. 5998164  
; GENERAL INFORMATION:  
; APPLICANT: LI, YI  
; APPLICANT: CAO, LIANG  
; APPLICANT: NI, JIAN  
; APPLICANT: GENTZ, REINER  
; APPLICANT: BULT, CAROL J.  
; APPLICANT: SUTTON III, GRANGER G.  
; TITLE OF INVENTION: Polynucleotides Encoding Human G-protein  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVE., NW, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,948A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04079  
; FILING DATE: 30-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEFFE, ERIC K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2540  
; TELEFAX: 202-371-2600  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 381 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide



Db 179 IMGWNCISALSSCSTVPLVHKHYILFCTVFTLLLSLVIILYCRYSLVTRSRRLTFR 238  
QY 241 AARRRARR-----LTKTYMLLAFVWGPFLGILLADVFSGNLMAOEYLRGMDWILA 294  
Db 239 KNISKASRSSENVALLTKTYIIVLSVFACWAPFLILLLDV-GCKVKTCDILFRAEYFLV 297  
QY 295 LAVNSAVNPITYSFRSREVCRAVLSFL-CCGC 326  
Db 298 LAVNSGNTPIITYTLTKEMKRAFRIRMSCCKC 330

## RESULT 13

PCT-US96-10618-4  
Sequence 4, Application PC/TUS9610618

GENERAL INFORMATION:  
APPLICANT: Coleman, Roger  
APPLICANT: Guegler, Karl J.  
APPLICANT: Au-Young, Janice  
APPLICANT: Bandman, Olga  
APPLICANT: Selbman, Jeffrey J.  
TITLE OF INVENTION: A NOVEL HUMAN EDG-2 RECEPTOR HOMOLOG  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/10618  
FILING DATE: 20-JUN-1996

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/000,352  
FILING DATE: 20-JUN-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/567,817

FILING DATE: 06-DEC-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Gaister, Debra J.

REGISTRATION NUMBER: 33,888

REFERENCE/DOCKET NUMBER: PF-0042 PCT

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 381 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 119130

PCT-US96-10618-4

Query Match 32.2%; Score 633; DB 5; Length 381;

Best Local Similarity 40.5%; Pred. No. 1e-50;

Matches 135; Conservative 63; Mismatches 125; Indels 10; Gaps 5;

QY 1 MNTGPPVAESCGQLAAGSHRLVLAHNSGRLAGRGSPEDGGGLGALRGLSVAASCLV 60  
Db 1 MGPSTPLVVAHRSVSDYVNIIDIVRHVYTKL-NISADKENSIKLISVFIILCCPT 59  
QY 61 VLENLVLAAITSHMSRRWVYCLVNTLSDLLTGAAYLANVLSGARTFRLAPAWFL 120

Db 60 ILENIFVLLTWTKKRRHMYFYIGNALSDLLAGVAATANLLSGATTYKLTLPQWFL 119  
QY 121 REGILFALAASTPSSLFTGERFATWVRVPAASGATKTSRVYGFGLCMLAALLGMP 180  
Db 120 REGSMFVALASVAFSLAIAIERITMLKMKLNG-SNNFRLLLSACWVLSLIGGLP 178  
QY 181 ILGNWCIAEDRCSSLPVSKRYIEFCVIFAGVATIMGIVGATFRLVOASGOKAPP 240  
Db 179 IMGWNCISALSSCSTVPLVHKHYILFCTVFTLLLSLVIILYCRYSLVTRSRRLTFR 238  
QY 241 AARRRARR-----LTKTYMLLAFVWGPFLGILLADVFSGNLMAOEYLRGMDWILA 294  
Db 239 KNISKASRSSENVALLTKTYIIVLSVFACWAPFLILLLDV-GCKVKTCDILFRAEYFLV 297  
QY 295 LAVNSAVNPITYSFRSREVCRAVLSFL-CCGC 326  
Db 298 LAVNSGNTPIITYTLTKEMKRAFRIRMSCCKC 330

## RESULT 14

US-08-118-270-73  
Sequence 73, Application US/08118270

Patent No. 5508384

GENERAL INFORMATION:

APPLICANT: Murphy, Randall B.  
APPLICANT: Schuster, David I.

TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN

NUMBER OF SEQUENCES: 348  
CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington  
STATE: D.C.  
COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/118,270

FILING DATE: 09-SEP-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236

FILING DATE: 10-SEP-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.

REGISTRATION NUMBER: 34,033

REFERENCE/DOCKET NUMBER: MURPHY-2A

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

TELEX: 248633

INFORMATION FOR SEQ ID NO: 73:

SEQUENCE CHARACTERISTICS:

LENGTH: 334 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-118-270-73

Query Match 30.5%; Score 600; DB 1; Length 334;

Best Local Similarity 44.2%; Pred. No. 9.9e-46;

Matches 122; Conservative 53; Mismatches 91; Indels 10; Gaps 4;

QY 58 CLVLENTLVLAATIRSHMSRRWVYCLVNTLSDLLTGAAYLANVLSGARTFRLAPAQ 117  
Db 11 CFTLENIFVLLTWTKKRRHMYFYIGNALSDLLAGVAATANLLSGATTYKLTLPQ 70





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: December 9, 2002, 12:03:06 ; Search time 21 Seconds

(without alignments)  
1757.887 Million cell updates/sec

Title: US-09-786-926-1

Perfect score: 1966

Sequence: 1 MNATGTPVAPESCOQLAAG.....RSLSPMRREPLSSISVSRT 384

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

## Database :

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 666   | 33.9        | 383    | 2     | 153870      |
| 2          | 656.5 | 33.4        | 362    | 2     | JC7559      |
| 3          | 645.5 | 32.8        | 378    | 2     | JC5245      |
| 4          | 633   | 32.2        | 381    | 2     | A35300      |
| 5          | 517.5 | 26.3        | 352    | 2     | JC1465      |
| 6          | 438   | 22.3        | 364    | 2     | JC5293      |
| 7          | 351   | 17.9        | 473    | 2     | A33117      |
| 8          | 344.5 | 17.5        | 180    | 2     | E48909      |
| 9          | 333   | 16.9        | 472    | 2     | S17595      |
| 10         | 299.5 | 15.2        | 330    | 2     | A55689      |
| 11         | 288.5 | 14.7        | 337    | 2     | S40454      |
| 12         | 285   | 14.5        | 347    | 2     | S70364      |
| 13         | 278.5 | 14.2        | 360    | 2     | S36750      |
| 14         | 248.5 | 12.6        | 362    | 2     | I65990      |
| 15         | 245   | 12.4        | 412    | 2     | D30341      |
| 16         | 245   | 12.4        | 363    | 2     | S48697      |
| 17         | 244.5 | 12.4        | 412    | 2     | A48978      |
| 18         | 244   | 12.4        | 409    | 2     | I48095      |
| 19         | 240.5 | 12.4        | 325    | 2     | JN0764      |
| 20         | 237.5 | 12.1        | 325    | 2     | JC5592      |
| 21         | 235   | 11.9        | 361    | 2     | B46137      |
| 22         | 233.5 | 11.9        | 410    | 2     | A48974      |
| 23         | 228.5 | 11.6        | 409    | 2     | S68247      |
| 24         | 228   | 11.6        | 572    | 2     | I39369      |
| 25         | 227.5 | 11.6        | 355    | 2     | S39028      |
| 26         | 222   | 11.3        | 372    | 2     | I49008      |
| 27         | 221.5 | 11.3        | 501    | 2     | JH0447      |
| 28         | 221   | 11.2        | 314    | 2     | S71420      |
| 29         | 221   | 11.2        | 314    | 2     | S70005      |

|    |       |      |     |   |        |                     |
|----|-------|------|-----|---|--------|---------------------|
| 30 | 220   | 11.2 | 387 | 2 | S55550 | 5-HT4s receptor -   |
| 31 | 220   | 11.2 | 406 | 2 | S55549 | serotonin 4 recept  |
| 32 | 216.5 | 11.0 | 317 | 2 | S45708 | MSH receptor - bov  |
| 33 | 215   | 10.9 | 325 | 2 | JC2193 | melanocortin recep  |
| 34 | 215   | 10.9 | 351 | 1 | A55962 | opsin, pineal glan  |
| 35 | 214   | 10.9 | 352 | 1 | I50081 | rhodopsin - green   |
| 36 | 213.5 | 10.9 | 348 | 1 | S23398 | rhodopsin - Chines  |
| 37 | 212.5 | 10.8 | 348 | 1 | JC4267 | opsin - rabbit      |
| 38 | 212   | 10.8 | 348 | 1 | QO80   | rhodopsin - bovine  |
| 39 | 211   | 10.7 | 560 | 2 | A38731 | alpha-1A adrenergic |
| 40 | 210.5 | 10.7 | 320 | 2 | S17177 | probable G protein  |
| 41 | 208   | 10.6 | 332 | 2 | I48933 | adenosine receptor  |
| 42 | 207.5 | 10.6 | 332 | 2 | A57055 | melanocortin recep  |
| 43 | 207.5 | 10.6 | 348 | 1 | S51677 | rhodopsin - rat     |
| 44 | 207.5 | 10.6 | 348 | 1 | QO8H   | rhodopsin - sheep   |
| 45 | 207   | 10.5 | 332 | 2 | JC1229 | adenosine receptor  |

## ALIGNMENTS

## RESULT 1

Eg-1 Orphan receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 24-Sep-1999

C:Accession: I53870

R:Lado, D.C.; Browe, C.S.; Gaskin, A.A.; Borden, J.M.; MacLennan, A.J.

Gene 149, 331-336, 1994

A:Title: Cloning of the rat eg-1 immediate-early gene: expression pattern suggests d

A:Reference number: I53870; M0ID:95047498; PMID:7959012

A:Accession: I53870

A&gt;Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: mRNA

A:Residues: 1-383 &lt;RES&gt;

A:Cross-references: EMBL:U010303; NID:G959396; PIDN:AAA83418.1; PID:G959397

C:Superfamily: G protein-coupled receptor eg-1

Query Match 33.9% Score 666; DB 2: Length 383;

Best Local Similarity 40.0%; Pred. No. 3e-50;

Matches 150; Conservative 69; Mismatches 126; Indels 30; Gaps 8;

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | MNATGTPVAPESCOQLAAGSHRLIVLHNSGRLAGRGSPEDGGLGALRLGSLVAASCLV | 60  |
| DB | 2   | VSSISIPVYKALRSQVEDYGNVDIVRHNYTGKL-NIGVEKHGIGKITSVVFILCCLI  | 60  |
| QY | 61  | VLENLVLIAITSHMRSRMYVYCLVNTLSDLTGAAYLANVLSGARTFRLAPQWFL     | 120 |
| DB | 61  | ILENIFVLLITFKTKKFRPMYVYIGNLALSDLAGVAYANLILSGATYKILPAQWFL   | 120 |
| QY | 121 | REGILFTALASTSLFTTGERFATVWRPVASGATKTSRVYGFGLCWLAAALGMP      | 180 |
| DB | 121 | REGSFEVALSASVSLALAIERTITMKMLHNG-SNSRSFLISACAVISILLAGLP     | 179 |
| QY | 181 | LIGNNLCAPDRCSSLLPYLSKRYITLFCVIFAGVATLITGLCAIRLVOASQ----    | 235 |
| DB | 180 | IMGNWNCISLSSCSSTVPLVYHKHILFCTVFTLLLSIYLXRIYSIVRTSRRLRFL    | 239 |
| QY | 236 | ---KAPPAARRKARILKTVMLILALVYCWGPLEGLLADYFGSNLMAOELRMDW      | 291 |
| DB | 240 | KNISAAAR--SSEKSLALKTIVILSVFIACMPLEPILLLLDV-GCKARTCDILYAEY  | 296 |
| QY | 292 | ILALAVNSAVNPITVYSPRSRCRAVLSFL-CCGCLRLMGK-----PGDCLARAVE    | 343 |
| DB | 297 | FLVALVNSGTPNPIITVITTKEMRRARIRISCKCKENGDSAGKFRPIIPGMFES---  | 353 |
| QY | 344 | AHSGASTDSSSLRPR  | 358 |
| DB | 354 | -----SKSDNSHPQ   | 363 |

RESULT 2  
JC7559



Db 239 KNISAKSRSENNALLKTYIIVLSVFIACAPLFIILLIDV-GCKVKTCIDILFRAEFLV 297

QY 295 LAVINSAVNPRIYSFRSREVCRAVLSFL-CCGC 326

Db 298 LAVLNGSTNPRIYTLTKEMRAFRIRIMSCCK 330

## RESULT 5

JC1465

probable G protein-coupled receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jul-2000

C:Accession: JC1465

R:Okazaki, H.; Ishizaka, N.; Sakurai, T.; Kurokawa, K.; Goto, K.; Kumada, M.; Takawa, Y. Blochem. Biophys. Res. Commun. 190, 1104-1109, 1993

A:Title: Molecular cloning of a novel putative G protein-coupled receptor expressed in H

A:Reference number: JC1465; MUID:93176155; PMID:8382486

A:Accession: JC1465

A:Molecule type: mRNA

A:Residues: 1-352 <OKA>

A:Cross-references: GB:AB016931; NID:g3445557; PIDN:BA32454.1; PID:g3445558

A:Experimental source: aortic smooth muscle

A:Superfamily: G protein-coupled receptor edg-1

C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot

F:67-95/Domain: transmembrane #status predicted <TM1>

F:110-128/Domain: transmembrane #status predicted <TM3>

F:148-173/Domain: transmembrane #status predicted <TM4>

F:190-210/Domain: transmembrane #status predicted <TM5>

F:234-255/Domain: transmembrane #status predicted <TM6>

F:272-293/Domain: transmembrane #status predicted <TM7>

F:13/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:142,145,218,219,329,330,331,332/Binding site: phosphate (Ser) (covalent) #status predi

F:313/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 26.3%; Score 517.5; DB 2; Length 352;

Best Local Similarity 38.4%; Pred. No. 2, 3e-37;

Matches 127; Conservative 50; Mismatches 129; Indels 25; Gaps 7;

QY 10 PESCOQLAAGSHRLIVLHNHSGRLAGRGPEDEGLGALRGSLVAAACLVLENLLVLA 69

Db 11 PEKVOE-----HYNYTKETLDM--QETPSRKVASAFIILICCAIVENLLVLI 56

QY 70 AITSHMSRRMYCYCLVNTITSDLLTGAAYLANVLSGARFPRLAPQWFLREGILFTAL 129

Db 57 AAVANSKFSHSMYFLGLNLAASDLLAGVAFAVANTLISGPVTLSTPLQMFARREGSAFTTL 116

QY 130 AASTFSLFTAGEFPATVVRPAESGATKTSRVYGFITGLCWLIALAALGLMPLLGWNCICA 189

Db 117 SASVFSLLAIAIEROVAIAK-VKLYGDKSCRMMLIGASWLISLILGGLPIIGWNCIDH 175

QY 190 FDRCSLLPLYSKRYTILFCVIFPAGVLTATMGILGAIFRLYOASGOKAPPAARRKRRRL 249

Db 176 LEACSTVPLPYAKHYVLCVVTIFSVIILAIYALYRIVFYVRRSHADVAGP---QPLAL 231

QY 250 LKTYVMIILALVCGMPLFGILLADVGSNMAQEVYLRGMOWIILALVINSVNPRIYSF 309

Db 232 LKTYVIVIGVFTICWLPKFSITLLD-STCPYRACPVLYKHAHYFAFATLSLNPVITYW 290

QY 310 RSREVCRAVLSFLCC---GCLRLGMRG--PG 335

Db 291 RSRDLRRVLRPLLCWRGKATGRGSGNP 321

## RESULT 6

JC5293

lysophosphatidic acid receptor - human

N:Alternate names: Edg2 protein

C:Species: Homo sapiens (man)

C:Date: 16-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 10-Dec-1999

C:Accession: JC5293

R:An, S.; Dickens, M.A.; Bleu, T.; Hallmark, O.G.; Goetzl, E.J.

Biochem. Biophys. Res. Commun. 231, 619-622, 1997

A:Title: Molecular cloning of the human Edg2 protein and its identification as a func

A:Reference number: JC5293; MUID:97224397; PMID:9070858

A:Contents: lung

A:Accession: JC5293

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-364 <ANA>

A:Cross-references: GB:U80811; NID:g1857424; PIDN:AAC51139.1; PID:g1857425

C:Superfamily: G protein-coupled receptor edg-1

Query Match 22.3%; Score 438; DB 2; Length 364;

Best Local Similarity 33.1%; Pred. No. 1, 9e-30;

Matches 111; Conservative 54; Mismatches 152; Indels 18; Gaps 6;

QY 1 MNATPTAPPSCOQLAAG-----HSRLIVLHNHSGR-LAGRGPEDEGLALRGISV 54

Db 1 MAISTSPISVISOPEFTMNEPQCFYNSIAFEYFRSKHLATMTWNV--KLYMGLGI 57

QY 55 AASCLVLENLLVLAITSNHSRRMYCYCLVNTITSDLLTGAAYLANVLSGARTFLA 114

Db 58 TYCITIMLANLLMYVAIVNRRFHPPIYLMANLAADFFAGLAFYLMFTGPNTRRLT 117

QY 115 PAQWFLREGILFTALASTFSLFTAGERFATVVRPAESGATKTSRYVGFICWILAA 174

Db 118 VSTWLLRQGLIDTSLTASVANILAIERHITVEFMQLHT-RMSNRRVYVIVYIMTAI 176

QY 175 ILGMLPLLGWNCICAFDDCSSLLPLYSKRYTILFCVIFPAGVLTATMGILGAIFRLYOASG 234

Db 177 VNGALPSVGNWTCIDIECNSSNAPILSYLVEFMAIFMLVFFVYVVAIHAIFGVQR 236

QY 235 QKAPPAARRKRR-----LKTVMILALVLCWGPLFGILLADVGSNMAQEVYLRGM 289

Db 237 MMSHRSSSPRRNRDPMMSLKTIVYVIGAFITCPTGVLLIDVCCPQCDVLAVER-- 294

QY 290 DWILALVINSVNPRIYSFRSREVCRAVLSFLCC 324

Db 295 FFLILAEPNSAMNPITYSYRDEKMSATFRQLIC 328

## RESULT 7

A33117

cannabinoid receptor CBI - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 05-Nov-1999

A:Accession: A33117; C55879

R:Katsuda, I.A.; Lolait, S.J.; Brownstein, M.J.; Young, A.C.; Bonner, T.I.

Nature 346, 561-564, 1990

A:Title: Structure of a cannabinoid receptor and functional expression of the cloned

A:Reference number: A33117; MUID:90332039; PMID:2165569

A:Accession: A33117

A:Molecule type: mRNA

A:Residues: 1-473 <MAT>

A:Cross-references: GB:X55812; NID:g1552375; PIDN:CAA3932.1; PID:g57249

R:Shire, D.; Carillon, C.; Keshnd, M.; Calandra, B.; Rinaldi-carmona, M.; Le Fur, G.

J. Biol. Chem. 270, 3726-3731, 1995

A:Title: An amino-terminal variant of the central cannabinoid receptor resulting from

A:Reference number: A55879; MUID:95181329; PMID:7876112

A:Accession: C55879

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-107 <SHD>

C:Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 17.9%; Score 351; DB 2; Length 473;

Best Local Similarity 30.1%; Pred. No. 9, 1e-23;

Matches 105; Conservative 61; Mismatches 129; Indels 54; Gaps 11;

QY 48 ALRGLSYAASCLVLENLLVLAITSNHSRR--VYICLVNTITSDLLTGAAYLANVLL 105

Db 119 AIAVLSLTGLTFVLENLLVLCVIL-HSRSLRCRPSYHIGSLAVADLGSVIFYVSFD 177

QY 106 SGARTFLAQAQWFLRGLLFTLAASTPSLFTAGERFATVVRPA-ESGATKTSRYVG 164

Db 178 FHFVHRKDSNVFLFKIGVTAFTASTASVGSFLTAIDRYSIHRPLAKRIVTPRKAYVA 237  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-472 <GER2>  
A:Cross-references: EMBL:X54937; NID:929914; PIDN:CAA38699.1; PID:929915  
Db 238 FC-LMWTIAIVAVLPPLGNCKKLOSCVDIFPFLIDENTYLMWIGVTSVLLFIYAYVA 296  
A:Title: An amino-terminal variant of the central cannabinoid receptor resulting from  
A:Reference number: A55879; MUID:95181329; PMID:7876112  
A:Accession: A55879  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-124 <SH1>  
A:Cross-references: GB:X81120  
C:Genetics:  
A:Gene: GDB:CNRL; CNR  
A:Cross-references: GDB:127354; OMIM:114610  
A:Map position: 6q14-q415  
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein  
F:117-142/Domain: transmembrane #status predicted <TM1>  
F:155-175/Domain: transmembrane #status predicted <TM2>  
F:188-212/Domain: transmembrane #status predicted <TM3>  
F:233-256/Domain: transmembrane #status predicted <TM4>  
F:275-299/Domain: transmembrane #status predicted <TM5>  
F:345-365/Domain: transmembrane #status predicted <TM6>  
F:378-399/Domain: transmembrane #status predicted <TM7>  
F:77,83/Binding site: carbohydrate (asn) (covalent) #status predicted  
Query Match 16.9%; Score 333; DB 2; Length 472;  
Best Local Similarity 29.4%; Pred. No. 3,3e-21;  
Matches 105; Conservative 56; Mismatches 116; Indels 80; Gaps 14;  
Db 407 HARBSMFPSCGTAQPLDMSMGSDCL-----HKHANNATSMHRAAES 449  
RESULT 8  
G protein-coupled receptor Gprcrl3 - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 01-Dec-1995 #sequence,revision 01-Dec-1995 #text\_change 19-May-2000  
C:Accession: E48909  
R:Miller, T.M.; Chen, Y.; Gilbert, D.J.; Moore, K.J.; Yu, L.; Simon, M.I.; Copeland, N.G.  
Genomics 18, 175-184, 1993  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 81 VYVCVNTITSLDGLGAATLAVNLISGARTFRLPAQWFLREGFLFTALASTSLTETA 140  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 7 MFLFGLNLAASDLGAVFANTLISGAVTSLTPVQWFAFVSAFTLSASVSLAIA 66  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 141 GERFATMPVAESGATKTSRYGFIQCLWLAALGMLPLGNMCLCAFRCSLLPLY 200  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 67 IEROVALAK-VKLYGSDKSRMLMIGASWILSLIGSLGNCKLKLKACSTVFLPY 125  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 201 SKRYTLFCVITFAGVLAITMGLYGAIFRLVQASGQKAPRPAARRKARRLKTVMLT 257  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 126 AKRYVLCVVTTFVILLAIVALYRIYVSSHADVAGP---QTALLKTVITVL 178  
RESULT 9  
cannabinoid receptor CB1 - human  
C:Species: Homo sapiens (man)  
C:Date: 03-Feb-1994 #sequence,revision 03-Feb-1994 #text\_change 21-Jul-2000  
C:Accession: S17595; S13668; A55879  
R:Gerard, C.M.; Mollereau, C.; Vassart, G.; Parmentier, M.  
Biochem. J. 279, 129-134, 1991  
A:Title: Molecular cloning of a human cannabinoid receptor which is also expressed in the  
A:Reference number: S17595; MUID:92028798; PMID:1718258  
A:Accession: S17595  
A:Molecule type: mRNA  
A:Residues: 1-472 <GER1>  
A:Cross-references: EMBL:X54937; NID:929914; PIDN:CAA38699.1; PID:929915  
R:Gerard, C.; Mollereau, C.; Vassart, G.; Parmentier, M.  
Nucleic Acids Res. 18, 7142, 1990  
A:Title: Nucleotide sequence of a human cannabinoid receptor cDNA.  
A:Reference number: S13668; MUID:91088303; PMID:2263478  
A:Accession: S13668

A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-472 <GER2>  
A:Cross-references: EMBL:X54937; NID:929914; PIDN:CAA38699.1; PID:929915  
R:Shire, D.; Carillon, C.; Kaghad, M.; Calandra, B.; Rinaldi-Carmona, M.; Le Fur, G.;  
J. Biol. Chem. 270, 3726-3731, 1995  
A:Title: An amino-terminal variant of the central cannabinoid receptor resulting from  
A:Reference number: A55879; MUID:95181329; PMID:7876112  
A:Accession: A55879  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-124 <SH1>  
A:Cross-references: GB:X81120  
C:Genetics:  
A:Gene: GDB:CNRL; CNR  
A:Cross-references: GDB:127354; OMIM:114610  
A:Map position: 6q14-q415  
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein  
F:117-142/Domain: transmembrane #status predicted <TM1>  
F:155-175/Domain: transmembrane #status predicted <TM2>  
F:188-212/Domain: transmembrane #status predicted <TM3>  
F:233-256/Domain: transmembrane #status predicted <TM4>  
F:275-299/Domain: transmembrane #status predicted <TM5>  
F:345-365/Domain: transmembrane #status predicted <TM6>  
F:378-399/Domain: transmembrane #status predicted <TM7>  
F:77,83/Binding site: carbohydrate (asn) (covalent) #status predicted  
Query Match 16.9%; Score 333; DB 2; Length 472;  
Best Local Similarity 29.4%; Pred. No. 3,3e-21;  
Matches 105; Conservative 56; Mismatches 116; Indels 80; Gaps 14;  
Db 48 ALRGVSAASCLVYLEMLVLAATSHMRSRW--VYCVNTITSLDGLGAATL----- 100  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 118 AIAVLSLTGLFTYLEMLVLCVTL-HSKSLRCPSPHFGLSADVLDSVLFVYSFID 176  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 101 -----ANVLISGARTFRLPAQWFLREGFLFTALASTSLFTAGERPATWRPY 151  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 177 FHFVHRKDSNVFLFKIGVTAFTASTASVGSFLTAIDRYSIHRPLAKRIVTPRKAYVA 237  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-472 <GER2>  
A:Cross-references: EMBL:X54937; NID:929914; PIDN:CAA38699.1; PID:929915  
R:Shire, D.; Carillon, C.; Kaghad, M.; Calandra, B.; Rinaldi-Carmona, M.; Le Fur, G.;  
J. Biol. Chem. 270, 3726-3731, 1995  
A:Title: An amino-terminal variant of the central cannabinoid receptor resulting from  
A:Reference number: A55879; MUID:95181329; PMID:7876112  
A:Accession: A55879  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-124 <SH1>  
A:Cross-references: GB:X81120  
C:Genetics:  
A:Gene: GDB:CNRL; CNR  
A:Cross-references: GDB:127354; OMIM:114610  
A:Map position: 6q14-q415  
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein  
F:117-142/Domain: transmembrane #status predicted <TM1>  
F:155-175/Domain: transmembrane #status predicted <TM2>  
F:188-212/Domain: transmembrane #status predicted <TM3>  
F:233-256/Domain: transmembrane #status predicted <TM4>  
F:275-299/Domain: transmembrane #status predicted <TM5>  
F:345-365/Domain: transmembrane #status predicted <TM6>  
F:378-399/Domain: transmembrane #status predicted <TM7>  
F:77,83/Binding site: carbohydrate (asn) (covalent) #status predicted  
Query Match 16.9%; Score 333; DB 2; Length 472;  
Best Local Similarity 29.4%; Pred. No. 3,3e-21;  
Matches 105; Conservative 56; Mismatches 116; Indels 80; Gaps 14;  
Db 152 A-ESGATKTSRYGFIQCLWLAALGMLPLGNMCLCAFRCSLLPLYSKRYLFCV 210  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 223 AKRYVLCVVTTFVILLAIVALYRIYVSSHADVAGP---QTALLKTVITVL 178  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 211 IFAGVLAITMGLYGAIFRLVQASGQKAPRPAARRKARRLKTVMLT 257  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 282 VTSVLLFPIYATMILMKAHSAVRAKIDRGTKSIIHTSEDKVQVTPDQARMDIRL 341  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 250 LKTVMLLAFLVQWGLFGLLDADVGSNLMQAEYLKGMWIL-----ALAVLSA 301  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 342 AKTVLTLVLLIICWGPLAIVYDVGK-----MNKLIKTVFAFCSMCLCINST 391  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 302 VNPIYSFRSREVCRAVLSFL--CCGCLRL--GMRPGDCLAR---AYEANSFAST 350  
A:Title: Identification, chromosomal location, and genome organization of mammalian G-pr  
A:Reference number: A48909; MUID:94116980; PMID:8288218  
A:Accession: E48909  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-180 <MUI>  
A:Cross-references: GB:L20334; NID:9438786; PIDN:AAA16846.1; PID:9438787  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor  
Query Match 17.5%; Score 344.5; DB 2; Length 180;  
Best Local Similarity 43.5%; Pred. No. 1.2e-22;  
Matches 77; Conservative 30; Mismatches 65; Indels 5; Gaps 2;  
Db 392 VNPIYALRSKDLRHAFRSMFPCSGTAQPLDMSMGSDCLHKHANNATSMHRAAES 449  
RESULT 10  
G protein-coupled receptor 3 - human  
C:Species: Homo sapiens (man)  
C:Date: 03-Mar-1995 #sequence,revision 03-Mar-1995 #text\_change 19-May-2000  
C:Accession: A55689; S58521; C55733  
R:Jismaa, T.P.; Kiefer, J.; Liu, M.L.; Baker, E.; Sutherland, G.R.; Shire, J.  
Genomics 24, 391-394, 1994  
A:Title: Isolation and chromosomal localization of a novel human G-protein-coupled re  
A:Reference number: A55689; MUID:95213036; PMID:7698767  
A:Accession: A55689  
A:Status: preliminary  
A:Molecule type: DNA; mRNA  
A:Residues: 1-330 <LIS>  
A:Cross-references: GB:L32830; GB:L32831; NID:9602311; PIDN:AAA75560.1; PID:9602312  
R:Eggerlickx, D.; Deneff, J.F.; Labbe, O.; Hayashi, Y.; Refetoff, S.; Vassart, G.; Parm



R:Munro, S.; Thomas, K.L.; Abu-Shaar, M.  
Nature 365, 61-65, 1993  
A:Title: Molecular characterization of a peripheral receptor for cannabinoids.  
A:Reference number: S36750; MUID:93368659; PMID:7689702  
A:Accession: S36750  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-360 <MUN>  
A:Cross-references: EMBL:X74328  
C:Superfamily: melanocortin receptor

Query Match 14.2%; Score 278.5; DB 2; Length 360;  
Best Local Similarity 28.8%; Pred. No. 1.4e-16;  
Matches 92; Conservative 57; Mismatches 138; Indels 33; Gaps 11;

QY 53 SVASAC-----LVLENLVLAITSHMSRRWYVCVNTTSLDGLGAAYLANVLS- 106  
DB 35 AVAVLCTLLGLSLALENVAVLVLLSSHQLRKPSYLEIG-----SLAGADFLASVVEAC 89  
QY 107 GATFRL-----APQMFIREGLFTTALAASFSLFTFAGRRFATVVRPAESGATKTSR 161  
DB 90 SFVNFHFHGVDSKAVFLFKTSVTMTFTASGSLTALDRLVCLRYPPSYKALLTRGR 149  
QY 162 VYGFICGLMLAALGLMGLMGLCAFDRCSSLLPYSKRYIFCLVIFAGVLAITMG 221  
DB 150 GLVTIGIMVLSALVSYLPIMKWTG-CF-RPCSELFRLPMDYLSMLLTAFLPSGIY 207  
QY 222 LYGAIF-----RLVQASG---QKAPRPAARRKARRLKTYMILAFVCGPLGLLLA 273  
DB 208 TYGHVLMKAKHQAHVASLGHODQVPGAMRMRIDVRLAKTGLVLAVALICMPFVLALMH 267  
QY 274 DVEGSLMAOEYLRCMDWILALAVNSAVNPITYSRSREVCRAVLSFLCCGCLRGMRG 333  
DB 268 SL-ATTLSDQYKKAFAFCSMCLINSMNVPIYALRSGETRS--AHHLAMKACVGR 323  
QY 334 PGCLARAVEAHSASTDS 353  
DB 324 LG---SEAKEEAPRRSVET 340

## RESULT 14

I65990  
G:protein-coupled receptor 6 - human  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 24-Sep-1999  
C:Accession: I65990  
R:Heiber, M.; Docherty, J.M.; Shah, G.; Nguyen, T.; Cheng, R.; Heng, H.H.Q.; Marchese, A.  
DNA Cell Biol. 14, 25-35, 1995  
A:Title: Isolation of three novel human genes encoding G protein-coupled receptors.  
A:Reference number: I53033; MUID:95134353; PMID:7832990  
A:Accession: I65990  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-362 <RES>  
A:Cross-references: GB:L36150; NID:9598156; PIDN:AAA63181.1; PID:9598157  
C:Genetics:  
A:Gene: GDB:GPR6  
A:Cross-references: GDB:371713; OMIM:600553  
A:Map position: 6q21-6q21  
C:Superfamily: melanocortin receptor

Query Match 12.6%; Score 248.5; DB 2; Length 362;  
Best Local Similarity 27.8%; Pred. No. 5.6e-14;  
Matches 93; Conservative 53; Mismatches 150; Indels 39; Gaps 11;

QY 10 PSCCOGLAAGSHRLIVLHNSGLAGR--GPEDEGLAGLGLSY--AASCLVLEN 64  
DB 38 PPAALALGAGGA-----NGSLELSQLSAGPPGLLPAVPMVDLLCVSGTVIAGN 90  
QY 65 LVLVAITSHMSRRWYVCVNTTSLDGLGAAYLANVLSGARFRLAPQW--FLRE 122  
DB 91 ALVVALIASTPALRTIPMEVIVGSLATADLLAGCLLIHAFQ-----YLVPSSTVSLIV 145

QY 123 GLTFTALASTFSLTTFAGRRFATVVRPAESGATKTSRYGFIQLCLLALLGMPLL 182  
DB 146 GFLVASPAASVSLIATYDRIYSLNALTYSRRTLGHHLLAAWTYSLGSLPVL 205  
QY 183 GWNCICAFDRCSLLPYSKRYIFCLVIFAGVLAITMGLYGAIFRLV-----QAS 233  
DB 206 GWNCIAERACSVVRPL-ARSHVALLSAFPMVFGIMLHVRLICQVVRHAHQIALQOH 264  
QY 234 GQAPRPAARRKARRLKTYMILAFVCGPLGLLLAFLVCGPLGLLLAFLVCGPLGLLLA 293  
DB 265 CLAPHLAATKRG---VGLTAVLVGTFGASWLP---FAIYCVGSHEDPAVITYA---TL 315  
QY 294 ALAVNSAVNPITYSRSREVCRAVLSFLCCGCLR 328  
DB 316 LPATNSMINDPIYAFRNOEIORA-LMLLCCCFQ 349

## RESULT 15

D30341  
G:protein-coupled receptor RDC8 - dog  
C:Species: Canis lupus familiaris (dog)  
C:Date: 07-Jun-1990 #sequence\_revision 30-Jun-1992 #text\_change 29-Sep-1999  
C:Accession: D30341; S12824  
R:Libert, F.; Parmentier, M.; Lefort, A.; Dinsart, C.; Van Sande, J.; Maenhaut, C.; S  
Science 244, 569-572, 1989  
A:Title: Selective amplification and cloning of four new members of the G protein-cou  
A:Reference number: A30341; MUID:89242119; PMID:2541503  
A:Accession: D30341  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-412 <LIB>  
A:Cross-references: EMBL:X14052; NID:9905; PIDN:CAA32210.1; PID:9906  
A:Note: The sequence given in Figure 2 is different from that shown in lacking 412-se  
R:Libert, F.; Parmentier, M.; Lefort, A.; Dumont, J.E.; Vassart, G.  
Nucleic Acids Res. 18, 1914, 1990  
A:Title: Complete nucleotide sequence of a putative G protein coupled receptor: RDC8.  
A:Reference number: S12824; MUID:90245608; PMID:21596628  
A:Accession: S12824  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-412 <LIB>

A:Cross-references: EMBL:X14052; NID:9905; PIDN:CAA32210.1; PID:9906  
C:Genetics:  
A:Gene: RDC8  
C:Superfamily: adenosine receptor A1  
C:Superfamily: G protein-coupled receptor; glycoprotein; transmembrane protein

Query Match 12.6%; Score 248.5; DB 2; Length 412;  
Best Local Similarity 27.1%; Pred. No. 6.4e-14;  
Matches 93; Conservative 54; Mismatches 141; Indels 55; Gaps 14;

QY 52 LSYAASCLVLENLVLAAITSHMSRRWYVCVNTTSLDGLG--AAVLANVLSGAR 109  
DB 12 VELALVLAITLGVNVCWAVMNSINQVTVFVSLAAADIAVGVALIPFATISRG-- 69  
QY 110 TFRILPAQWFLRGLFTALASTFSLTTFAGRRFATVVRPAESGATKTSRYGFIQL 169  
DB 70 -FCAACHNCLFACFVYLVLTQSSIFSLAIALDRIAIRPLRNGLVGTAKGIIVG 128  
QY 170 WLAALGLMLPLGWN-----CLCAFDRCSSLLPL-YSKRYIFCL 209  
DB 129 WLSFAIGLTPMLGWNCSSQPKRGNYSGCGEGVACLFPE--DVPVMVNYVNFAP 185  
QY 210 VIRAGVLAITMGLYGAIF-----RLVQASGQKAPRPAAR--RKARLLKTYMILAF 261  
DB 186 VLVV--LLMLGVYRIFLFAARQLKQESOPLPERRASTLQEVHAASLAIYGLFA 243  
QY 262 VCGPL-----FGLLADVGSNLMAOEYLRGMWILALAVNSAVNPITYSRSREVC 316  
DB 244 LCVLPHTIINCFEPCPECSHAPLWM-YL-----TIVLSHTNSVVPFIYAVRIREFRQ 297  
QY 317 AVLSFLCCGCLRLGMRG---GDCLARAVEAHSASTDSLR 356

Mon Dec 9 12:35:25 2002

us-09-786-926-1.rpr

Page 7

Db 298 TFRKIIRSHVLR--RREPPKAGTISARALAAH-GSDGEQISLR 337

Search completed: December 9, 2002, 12:05:00  
Job time : 22 secs

---





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OK protein - protein search, using sw model

Run on: December 9, 2002, 12:03:01 ; Search time 14 Seconds  
(without alignments)  
1137.636 Million cell updates/sec

Title: US-09-786-926-1

Perfect score: 1966  
Sequence: 1 MNATGTPVAPBSCQQLAAG.....RSLSPMRREPLSSISVRSI 384

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 666   | 33.9        | 383    | 1  | EDG1_RAT    |
| 2          | 645.5 | 32.8        | 378    | 1  | EDG3_HUMAN  |
| 3          | 644   | 32.8        | 382    | 1  | EDG1_MOUSE  |
| 4          | 633   | 32.2        | 381    | 1  | EDG1_HUMAN  |
| 5          | 517.5 | 26.3        | 352    | 1  | H218_RAT    |
| 6          | 444.5 | 22.6        | 364    | 1  | EDG2_MOUSE  |
| 7          | 439   | 22.3        | 393    | 1  | EDG2_SHEEP  |
| 8          | 438   | 22.3        | 364    | 1  | EDG2_HUMAN  |
| 9          | 432.5 | 22.0        | 364    | 1  | EDG2_BOVIN  |
| 10         | 351   | 17.9        | 473    | 1  | CB1R_MOUSE  |
| 11         | 351   | 17.9        | 473    | 1  | CB1R_RAT    |
| 12         | 346.5 | 17.6        | 470    | 1  | CB1B_FURGU  |
| 13         | 345   | 17.5        | 472    | 1  | CB1R_FELCA  |
| 14         | 344.5 | 17.5        | 180    | 1  | EDG1_MOUSE  |
| 15         | 333   | 16.9        | 472    | 1  | CB1R_HUMAN  |
| 16         | 330.5 | 16.8        | 473    | 1  | CB1R_POEUG  |
| 17         | 330   | 16.8        | 468    | 1  | CB1A_FURGU  |
| 18         | 317.5 | 16.1        | 473    | 1  | CB1R_TARGR  |
| 19         | 299.5 | 15.2        | 330    | 1  | CB1R_HUMAN  |
| 20         | 288.5 | 14.7        | 330    | 1  | CB1R_MOUSE  |
| 21         | 285   | 14.5        | 347    | 1  | CB2R_MOUSE  |
| 22         | 285   | 14.5        | 360    | 1  | CB2R_RAT    |
| 23         | 281.5 | 14.3        | 360    | 1  | CB2R_HUMAN  |
| 24         | 258.5 | 13.1        | 334    | 1  | GP12_HUMAN  |
| 25         | 256.5 | 13.0        | 334    | 1  | GP12_MOUSE  |
| 26         | 256.5 | 13.0        | 334    | 1  | GP12_RAT    |
| 27         | 248.5 | 12.6        | 362    | 1  | CB1R_HUMAN  |
| 28         | 248.5 | 12.6        | 412    | 1  | AA2A_CANFA  |
| 29         | 245   | 12.5        | 412    | 1  | AA2A_HUMAN  |
| 30         | 244.5 | 12.4        | 412    | 1  | AA2A_HUMAN  |
| 31         | 236.5 | 12.0        | 325    | 1  | MC5R_HUMAN  |
| 32         | 235   | 12.0        | 361    | 1  | OP5B_CHICK  |
| 33         | 234.5 | 11.9        | 325    | 1  | MC5R_PANTR  |

|    |       |      |     |   |            |        |              |
|----|-------|------|-----|---|------------|--------|--------------|
| 34 | 233.5 | 11.9 | 410 | 1 | AA2A_RAT   | P30543 | rattus norv  |
| 35 | 231.5 | 11.8 | 561 | 1 | ALAD_RAT   | P23944 | rattus norv  |
| 36 | 230   | 11.7 | 402 | 1 | OPN3_HUMAN | Q9H1Y3 | homo sapien  |
| 37 | 230   | 11.7 | 410 | 1 | AA2A_MOUSE | Q60613 | mus musculu  |
| 38 | 228.5 | 11.6 | 409 | 1 | AA2A_CAVPO | P46616 | cavia porce  |
| 39 | 228   | 11.6 | 572 | 1 | ALAD_HUMAN | P25100 | homo sapien  |
| 40 | 227.5 | 11.6 | 348 | 1 | OPSD_TURTR | O62798 | turslops tr  |
| 41 | 227.5 | 11.6 | 355 | 1 | OP5B_ASTFA | P51472 | astyanax fa  |
| 42 | 227.5 | 11.6 | 388 | 1 | 5H4_CAVPO  | O70528 | cavia porce  |
| 43 | 224.5 | 11.4 | 576 | 1 | 5H4_MOUSE  | P97288 | mus musculu  |
| 44 | 223   | 11.3 | 388 | 1 | ALAD_RABIT | O02666 | oryctolagus  |
| 45 | 222   | 11.3 | 314 | 1 | MSHR_CHICK | P55167 | gallus galli |

#### ALIGNMENTS

| RESULT 1 | EDG1_RAT   | STANDARD: | PRT: | 383 AA. |
|----------|--|-----------|------|---------|
| AC       | P48303;  |           |      |         |
| DT       | 01-FEB-1996 (Rel. 33, Created)   |           |      |         |
| DT       | 01-FEB-1996 (Rel. 33, Last sequence update)  |           |      |         |
| DT       | 15-JUL-1998 (Rel. 36, Last annotation update)  |           |      |         |
| DE       | Probable G protein-coupled receptor EDG-1.   |           |      |         |
| GN       | EDG1.  |           |      |         |
| OS       | Rattus norvegicus (Rat).   |           |      |         |
| OC       | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  |           |      |         |
| OC       | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  |           |      |         |
| OX       | NCBI_TaxID=10116;  |           |      |         |
| RN       | [1]  |           |      |         |
| RP       | SEQUENCE FROM N.A.   |           |      |         |
| RC       | TISSUE=Cerebellum;   |           |      |         |
| KX       | MEDLINE=95047496; PubMed=7959012;  |           |      |         |
| RA       | Lado D.C., Browe C.S., Gaskin A.A., Borden J.M., MacLennan A.J.;   |           |      |         |
| RT       | "Cloning of the rat edg-1 immediate-early gene: expression pattern suggests diverse functions."  |           |      |         |
| RL       | Gene 149:331-336(1994).  |           |      |         |
| CC       | -1- FUNCTION: IT IS POSSIBLE THAT IT ACTS AS A RECEPTOR FOR A TROPHIC FACTOR AND IS INVOLVED IN THE SURVIVAL OF BRAIN CELLS.   |           |      |         |
| CC       | -1- SUBCELLULAR LOCATION: Integral membrane protein.   |           |      |         |
| CC       | -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT EMBRYONIC DAY 15. AT POSTNATAL DAY 14 DETECTED IN SKIN, SPLEEN, LIVER, KIDNEY, HEART, TESTICLE, LUNG AND BRAIN. AT ADULTHOOD IS MOST ABUNDANT IN BRAIN.   |           |      |         |
| CC       | -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  |           |      |         |
| CC       | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ). |           |      |         |
| CC       | -----  |           |      |         |
| DR       | EMBL; U10303; AA83418.1; -   |           |      |         |
| DR       | InterPro: IPR000276; GPCR_Rhodpsn.   |           |      |         |
| DR       | InterPro: IPR004061; S1PReceptor.  |           |      |         |
| DR       | Pfam: PF00001; 7tm.1; 1.   |           |      |         |
| DR       | PRINTS; PR00523; GPCR_RHODPSN.   |           |      |         |
| DR       | PRINTS; PR01523; S1PRECEPTOR.  |           |      |         |
| DR       | PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.  |           |      |         |
| DR       | PROSITE; PS0262; G-PROTEIN_RECP_F1_2; 1.   |           |      |         |
| KW       | G-protein coupled receptor; Transmembrane; Glycoprotein; Phosphorylation; Lipoprotein; Palmitate.  |           |      |         |
| KW       | DOMAIN 1   |           |      |         |
| FT       | TRANSMM 48   |           |      |         |
| FT       | TRANSMM 73   |           |      |         |
| FT       | TRANSMM 79   |           |      |         |
| FT       | TRANSMM 80   |           |      |         |
| FT       | DOMAIN 109   |           |      |         |
| FT       | TRANSMM 123  |           |      |         |
| FT       | TRANSMM 142  |           |      |         |
| FT       | DOMAIN 161   |           |      |         |
| FT       | TRANSMM 187  |           |      |         |
| FT       | DOMAIN 202   |           |      |         |

```
FT TRANSMEM 203 223 5 (POTENTIAL).
FT DOMAIN 224 257 6 (POTENTIAL).
FT TRANSMEM 258 279 6 (POTENTIAL).
FT DOMAIN 280 295 6 (POTENTIAL).
FT TRANSMEM 296 316 7 (POTENTIAL).
FT DOMAIN 317 383 7 (POTENTIAL).
FT CARBOHYD 31 31 7 (POTENTIAL).
FT LIPID 329 329 7 (POTENTIAL).
FT MOD.RES 354 354 7 (POTENTIAL).
SQ SEQUENCE 383 AA: 42745 MW: 090BA0AE09D4F3 CRC64:

Query Match 33.9%; Score 666; DB 1; Length 383;
Best Local Similarity 40.0%; Pred. No. 2.1e-37;
Matches 150; Conservative 69; Mismatches 126; Indels 30; Gaps 8;

OY 1 MNATGPVAPESQOOLAGSHRLVLAHNSGRGLAGRGPDGSLGRLGSLVASCCLV 60
DB 2 VSTSLPVPVKALRSOVSDDYGNVDIIVRHNNGKL-NIGVEKDHGIKLTSVFILICLI 60
OY 61 VLENLVLVAITSHMSRRVYICLVNITLSDLTGAAYLANVLGARTFRLAPQWEL 120
DB 61 ILENFVLLITKTKKHFHPMYEFGNLSLGLAGVATNLLSGATTKLIPQWFL 120
OY 121 REGCLFTALASTFSLFTAGGERFATWVPAESGATTSRYGFIQLCWLLALGLMP 180
DB 121 RGSMEVALASVFSILAIERITMLKMLHNG-SNSSRSFLLISACWYISLIGLP 179
OY 181 LIGMCLAFDRCSLLPYSKRYTLFCLVIFAGVLAITMGLKGIFFLVQASGQ----- 235
DB 180 IGMWNCISLSCSTVLPHKHYTLFCTVFTLLLSIVLLCYISLVTRSRRLTFR 239
OY 236 ----KAPRPARRKARLLKTVLMLLAFVCMGPFLLADVFGSLMAQEYLRGMW 291
DB 240 KNISKASR--SEKSIALLKTYIVLVSVFIACWAPFLFLLLDV-GCAKAKCDILIKREY 296
OY 292 ILALVMSAVNPITTSRSRRCVCAVLSFL-CCGCLRLKRG-----PGDCLARAVE 343
DB 297 FLVLAVLNSGTPITLYTLTNEMRAFLRIISCKCPNGSAGKFKRPILIPMEFSR--- 353
OY 344 AHSGASTTDSLRPR 358
DB 354 ----SKSDNSHPQ 363

RESULT 2
EDG3_HUMAN STANDARD: PRT: 378 AA.
ID EDG3_HUMAN Q99500:
AC Q99500:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lysosphingolipid receptor (EDG-3).
GN EDG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97032811; PubMed=8878560;
RA Ymaguchi F., Tokuda M., Hatase O., Brenner S.;
RT "Molecular cloning of the novel human G protein-coupled receptor
RT (GPCR) gene mapped on chromosome 9.";
RL Biochem. Biophys. Res. Commun. 227:608-614(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98072391; PubMed=9409733;
RA An S., Bleu T., Huang W., Hallmark O.G., Coughlin S.R., Goetzl E.J.;
RT "Identification of cDNAs encoding two G protein-coupled receptors for
RT lysosphingolipids";
RL FEBS Lett. 417:279-282(1997).
```

```
CC -1- FUNCTION: ORPHAN RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES, BUT MOST ABUNDANTLY
CC IN HEART, PLACENTA, KIDNEY, AND LIVER.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.sib.ch).
CC -----
DR EMBL: X83864; CA58744.1; -.
DR EMBL: AF022139; AAC51906.1; -.
DR Genew: HGNC:3167; EDG3.
DR MIM: 601965; -.
DR InterPro: IPR004062; EDG3receptor.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR01524; EDG3RECEPTOR.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PRINTS: PR01523; S1PRECEPTOR.
DR PROSITE: PS00237; G_PROTEIN_RECPT_FL1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECPT_FL2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein.
KW DOMAIN 1 40
FT TRANSMEM 41 65
FT DOMAIN 66 72
FT TRANSMEM 73 101
FT DOMAIN 102 115
FT TRANSMEM 116 134
FT DOMAIN 135 153
FT TRANSMEM 154 179
FT DOMAIN 180 195
FT TRANSMEM 196 216
FT DOMAIN 217 243
FT TRANSMEM 244 265
FT DOMAIN 266 281
FT TRANSMEM 282 302
FT TRANSMEM 303 378
FT CARBOHYD 15 15
SQ SEQUENCE 378 AA: 42294 MW: 79A00306203F439F CRC64:

Query Match 32.8%; Score 645.5; DB 1; Length 378;
Best Local Similarity 42.2%; Pred. No. 4.8e-36;
Matches 147; Conservative 55; Mismatches 133; Indels 13; Gaps 6;

OY 28 HYHSGRLAGRGPDGGGALRGSLVASCCLVLENLVLVAITSHMSRRVYICLVN 87
DB 21 HYIYVKLALGRLEKASEGSTLTITVLELVICSTVLENLWLAIAIKNNKFNRMTEFGN 80
OY 88 ITLSDLLTGAAYLANVLGARTFRLAPQWELREGCLFTALASTFSLFAGGERFAT 147
DB 81 LALCDLLAGIAVKVNTLMGSKTFSLPTVWFLRGSNMVVALGASTCLALAIERHLM 140
OY 148 VAPVSGATKTSRYVGFGLCWLLAALGLPLIGMNCCLAFDRCSLLPYSKRYTLF 207
DB 141 IK-MRPYDANKRRHVFLLGCMCLIAFTGLALPIIGMNCILHMLPDCSTILPYSKYIAF 199
OY 208 CIVIFAGVLAITMGYGAIFRLVOASGOKAPRPARRKARLLKTVLMLLAFVCM 267
DB 200 CISIFAILVTVIILKARYIFLVKSSRRVANHNSRSMALLKTVIVVSVFTACWSP 259
OY 268 FGLLADVFGSLMAQEYLRGMWTLALAVLNSAVNPITTSRSRRCVCAVLSFLCC 327
DB 260 FILFLIDV-ACRYOACPLIFKQWFLVLAIVNSANNPIYTLASGEMRAFR-LVNC 317
OY 328 RLGMSGPGCLARAYVAHSGASTTDSLRPRDSFGSGLSRMRKPL 375
DB 318 ---VRGRG---ARASPIOPALDPS----RSKSSSSSNSSHSPKVEDL 355
```

```

RESULT 3
EDGL_MOUSE
ID EDGL_MOUSE STANDARD: PRT: 382 AA.
AC 008530;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Probable G protein-coupled receptor EDG-1.
GN EDG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRATN-BALB/c; TISSUE=Liver;
RX MEDLINE=97369927; PubMed=9226368;
RA Liu C.H., Hla T.;
RT "The mouse gene for the inducible G-protein-coupled receptor edg-1.";
RL Genomics 43:15-24(1997).
CC -1- FUNCTION: THIS INDUCIBLE EPITHELIAL CELL G-PROTEIN-COUPLED
CC RECEPTOR MAY BE INVOLVED IN THE PROCESSES THAT REGULATE THE
CC DIFFERENTIATION OF ENDOTHELIAL CELLS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES
CC WITH HIGHEST LEVELS IN BRAIN, SPLEEN AND HEART. LOWER LEVELS
CC FOUND IN LUNG, PLACENTA, MUSCLE, LIVER, UTERUS AND KIDNEY. VERY
CC LOW LEVELS IN TESTIS AND THYMUS.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: U04811; AAC53294.1; -.
DR MGD: MGI:1096355; Edg1.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR004061; S1Preceptor.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHDOPSN.
DR PRINTS: PR01523; S1PRECEPTOR.
DR PROSITE: PS00237; G_PROTEIN_RECPT_F1.1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECPT_F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 46 EXTRACELLULAR (POTENTIAL).
FT TRANSHEM 47 71 1 (POTENTIAL).
FT TRANSHEM 72 78 CYTOPLASMIC (POTENTIAL).
FT TRANSHEM 79 107 2 (POTENTIAL).
FT DOMAIN 108 121 EXTRACELLULAR (POTENTIAL).
FT TRANSHEM 122 140 3 (POTENTIAL).
FT DOMAIN 141 159 CYTOPLASMIC (POTENTIAL).
FT TRANSHEM 160 185 4 (POTENTIAL).
FT DOMAIN 186 201 EXTRACELLULAR (POTENTIAL).
FT TRANSHEM 202 222 5 (POTENTIAL).
FT TRANSHEM 223 256 CYTOPLASMIC (POTENTIAL).
FT TRANSHEM 257 278 6 (POTENTIAL).
FT DOMAIN 279 294 EXTRACELLULAR (POTENTIAL).
FT TRANSHEM 295 315 7 (POTENTIAL).
FT DOMAIN 316 382 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 30 30 N-LINKED (GLYCAC. ...) (POTENTIAL).
FT LIPID 328 328 PALMITATE (BY SIMILARITY).
FT MOD_RES 353 353 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 382 AA; 42571 MW; A004893491C8957 CRC64;

```

```

Query Match 32.8%; Score 644; DB 1; Length 382;
Best Local Similarity 38.9%; Pred. No. 6; le-36;

Matches 145; Conservative 67; Mismatches 135; Indels 26; Gaps 7;
QY 1 MNATGTPVAPESCOQLAAGHSRLIVLHYNHSGRIAGRGCPEDGLALGLSVAACTV 60
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MVSSTIPEVVALSSVSDYGVNDIIVRHNYTGKL-NIGAKRKDGIKLTSVFTLCCFI 59
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 VLENLVLAAITSHMSRWVYICLVNTITSLTGAAYLANVLSARFRLAPQWFL 120
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 60 ILENIFVLITWTKFENRPMYFIGNLALSLDLAGAAYANLLSGATYKLTLPQWFL 119
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 REGLEFALAASTFSLTFAGREFATWVRPAVESGATKTSRVYFGLCWLALALGMP 180
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 120 REGSMFVALSAVSFLALAIERYITMLKMLHNG-SNSSFLLISACVVISLIIIGLP 178
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 LIGMNCICADRCSSLLPYRSKRYILFCLVIFAGVLAITIMLYAIPRVOASGK- 236
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 179 SMGNCTSSLSGCTVLPYRKHVILFCTVFTLLSLALYCRISLVTRSRRLTER 238
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 --APRPAARKARRLKTLVLMILAFVCGPFLGLADVFSGNLMAOYLKGMWIL 293
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 239 KNISGSRSSSEKSLALFTVYIVLSVFIACWAPLFIILLIDV-GCKAKTCDILYKAEVFL 297
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 ALAVINAVNPITYSRSEKRAVLSFL-CCGCLRLGMRG-----PDCLARAVEAH 345
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 298 VLAIVNSGTPILYITLTKEMRAFRIRVSCCKPMDSGKFRPIIPMEFSR----- 352
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 346 SGASTTDSILRPR 358
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 353 ---SKDINSHPQ 362
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
EDGL_HUMAN
ID EDGL_HUMAN STANDARD: PRT: 381 AA.
AC P21453;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable G protein-coupled receptor EDG-1.
GN EDG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Umbilical vein endothelial cells;
RX MEDLINE=90264425; PubMed=2160972;
RA Hla T., Maciag T.;
RT "An abundant transcript induced in differentiating human endothelial
RT cells encodes a polypeptide with structural similarities to
RT G-protein-coupled receptors."
RL J. Biol. Chem. 265:9308-9313(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=98072391; PubMed=9409733;
RA An S., Bleu T., Huang W., Hallmark O.G., Coughlin S.R., Goetzl E.J.;
RT "Identification of cDNAs encoding two G protein-coupled receptors for
RT lysophingolipids."
RL FEBS Lett. 417:279-282(1997).
CC -1- FUNCTION: THIS INDUCIBLE EPITHELIAL CELL G-PROTEIN-COUPLED
CC RECEPTOR MAY BE INVOLVED IN THE PROCESSES THAT REGULATE THE
CC DIFFERENTIATION OF ENDOTHELIAL CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS, AND TO A LESSER EXTENT, IN
CC VASCULAR SMOOTH MUSCLE CELLS, FIBROBLASTS, MELANOCYTES, AND CELLS
CC OF EPITHELIOID ORIGIN.
CC -1- INDUCTION: BY THE TUMOR PROMOTER PHORBOL 12-MYRISTATE 13-ACETATE
CC (PME) IN THE PRESENCE OF CYCLOHEXIMIDE.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M31210; AAA52336.1; -  
 DR EMBL: AF022137; AAC51905.1; -  
 DR PIR: A35300; A35300.  
 DR Genew: HGNC:3165; EDG1.  
 DR MIM: 601974; -  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR InterPro: IPR004061; S1PRECEPTOR.  
 DR Pfam: PF00001; 7tm.1; 1.  
 DR PRINTS: PR00237; GPCR\_RHODOPSIN.  
 DR PRINTS: PR01523; S1PRECEPTOR.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECPT-FL\_1; 1.  
 DR PROSITE: PS00262; G\_PROTEIN\_RECPT-FL\_2; 1.  
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Phosphorylation; Lipoprotein; Palmitate.  
 FT DOMAIN 1 46  
 FT TRANSMEM 47 71  
 FT DOMAIN 72 78  
 FT TRANSMEM 79 107  
 FT DOMAIN 108 121  
 FT TRANSMEM 122 140  
 FT DOMAIN 141 159  
 FT TRANSMEM 160 185  
 FT DOMAIN 186 201  
 FT TRANSMEM 202 222  
 FT DOMAIN 223 255  
 FT TRANSMEM 256 277  
 FT DOMAIN 278 293  
 FT TRANSMEM 294 314  
 FT DOMAIN 315 381  
 FT CARBOHYD 30 30  
 FT CARBOHYD 36 36  
 FT LIPID 327 327  
 FT MOD\_RES 352 352  
 FT SEQUENCE 381 AA; 42695 MW; D706DA4C74C2ED93 CRC64;

Query Match 32.2%; Score 633; DB 1; Length 381;  
 Best Local Similarity 40.5%; Pred. NO. 3.2e-35;  
 Matches 135; Conservative 63; Mismatches 123; Indels 10; Gaps 5;

QY 1 MNAETPVAPESCOQLAAGSHRLIVLYNHSGLRAGRGEGDGLALGSLVAASCLV 60  
 DB 1 MGPTSVPLKVAHRSVSDVNVNIVRYNTGKL-MISADKENSILKTSVFLICPFI 59  
 QY 61 VLENLVLVAITSHMSRMVYVCVNTTSLDLTGAAYLANVLSGARFRLAPQWFL 120  
 DB 60 ILENIVLITWTKKFFHPPMYTFIGNALSDLLAGVAYRANLLISATYIKLTPQWFL 119  
 QY 121 REGILFTALAASTFSLFTAGREFAVYRPAVESGATKTSRYVGIGLMLLAALGMDP 180  
 DB 120 REGSMFVALASAVSFLLAIAIRKITYMKMLHNG-SINFEFLFLISACWVSLILGGDP 178  
 QY 181 ILGMNCLCAFDRCSLLPLYSKRYLFLFVAGLAFATIMGLYALPRLVQASGOKARPP 240  
 DB 179 IMGMNCLISLSSCYLPLRYHKHYFLFCTYVFTLLLSIYLIKRYISLVRTSRRLTR 238  
 QY 241 AARRKARR-----LKTVMMLLFLVYCGPRLFGLLADVFGSNLMAQETLRGMDWILA 294  
 DB 239 KNISKSRSENVALLKTYIVLVSFIACMAPELILLLLDV-GKVKYTCIDILFRAYFLV 297  
 QY 295 LAVLNAVNPITYSFRSRECAVLSFL-CCGC 326  
 DB 298 LAVLNGSTNPITITLLNKMRAFLRIMSCCKC 330

RESULT 5  
 H218\_RAT

ID H218\_RAT STANDARD: PRT: 352 AA.  
 AC P47752;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Probable G protein-coupled receptor H218 (AGRL6) (lysosphingolipid  
 DE receptor).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteleia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain, and Lung;  
 RX MEDLINE=94373324; PubMed=8087418;  
 RA MacLennan A.J., Browe C.S., Gaskin A.A., Lado D.C., Shaw G.;  
 RT "Cloning and characterization of a putative G-protein coupled  
 RT receptor potentially involved in development."; Mol.  
 RL Mol. Cell. Neurosci. 5:201-209(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=ortic smooth muscle;  
 RX MEDLINE=93176155; PubMed=8382486;  
 RA Okazaki H., Ishizaka N., Sakurai T., Kurokawa K., Goto K.;  
 RT "Molecular cloning of a novel putative G protein-coupled receptor  
 RT expressed in the cardiovascular system."; Biochem. Biophys. Res. Commun. 190:1104-1109(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98072991; PubMed=9409733;  
 RA An S., Blev T., Huang W., Hallmark O.G., Coughlin S.R., Goetzl E.J.;  
 RT "Identification of cDNAs encoding two G protein-coupled receptors for  
 RT lysophingolipids." FEBS Lett. 417:279-282(1997).  
 RL [1]  
 RT FUNCTION: POSSIBLE RECEPTOR FOR THE LYSOPHINGOLIPIDS SPHINGOSINE  
 CC 1-PHOSPHATE (SLP) AND DIHYDRO-SLP.  
 CC 1- SUBCELLULAR LOCATION: INTEGRAL IN ALL DEVELOPING TISSUES WITH  
 CC 1- TISSUE SPECIFICITY: EXPRESSED IN PRIMITIVE, TRANSFORMED CELLS. RELATIVE  
 CC HIGHEST LEVELS DETECTED IN PRIMITIVE, TRANSFORMED CELLS. RELATIVE  
 CC ABUNDANCE: LUNG > KIDNEY = SKIN = GUT > SPLEEN > BRAIN > LIVER.  
 CC 1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U10699; AAA19241.1; -  
 DR EMBL: AF022138; AAC33494.1; -  
 DR InterPro: IPR004063; EDG5receptor.  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR InterPro: IPR004061; S1PRECEPTOR.  
 DR Pfam: PF00001; 7tm.1; 1.  
 DR PRINTS: PR01525; EDG5RECEPTOR.  
 DR PRINTS: PR00237; GPCR\_RHODOPSIN.  
 DR PRINTS: PR01523; S1PRECEPTOR.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECPT-FL\_1; FALSE\_NG.  
 DR PROSITE: PS00262; G\_PROTEIN\_RECPT-FL\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;  
 KW Palmitate.  
 FT DOMAIN 1 34  
 FT TRANSMEM 35 59  
 FT DOMAIN 60 66  
 FT TRANSMEM 67 95  
 FT DOMAIN 96 109  
 FT TRANSMEM 110 128  
 FT DOMAIN 129 147  
 FT TRANSMEM 148 173

```

FT DOMAIN 174 189 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 190 210 5 (POTENTIAL).
FT DOMAIN 211 233 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 234 253 6 (POTENTIAL).
FT DOMAIN 256 271 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 272 292 7 (POTENTIAL).
FT DOMAIN 293 352 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 19 19 N-LINKED (GLCNAC...) (POTENTIAL).
FT LIPID 305 305 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 352 AA; 38734 MW; 9C933A18E756CELE CRC64;

Query Match 26.3%; Score 517.5; DB 1; Length 352;
Best Local Similarity 38.4%; Pred. No. 1,36-27;
Matches 127; Conservative 50; Mismatches 129; Indels 25; Gaps 7;

OY 10 PSCCOOLAAGHSHRLVLYHNSGRLAGRGEPEDGGLGALGLSVASCLVLENLVLVA 69
DB 11 PEKVOE-----HNYTKETLDM--QETPSRKVASAFITIIICCAIYVENLVLVI 56
OY 70 AITSMKSRKRWYCYLVNTITSLDITGAAYLANVLSGARFTRLPAPQWFLREGILFTAL 129
DB 57 AAVNRSKFSHMYTLFGNLAAASDLAAGAAEVANLTLISGPVTLISLPQWFAREGSAFTTL 116
OY 130 AASTSLFTAGEFRATVPRVAESGATRTSRVYFGICMLLALLGLMPLLGNCICA 189
DB 117 SASVSLSLAIEROYAIAK-VKLYGSDKSCMLMIGASMLISITLIGPLIGNCIDH 175
OY 190 FDRGSSLLPLYSKRYTLFCLVTFAGVATIMGLYALIFRVOASGOKAPRPAARRKRL 249
DB 176 LEAGSTVLEPLAKHYLVCTVFESVILLALVALYRIVFYVSSHADVAGP---QTLAL 231
OY 250 LRTVIMLILAFVCGPLFGILLADYFGSNLMAOETLGMOMLIALVINSVNPITISF 309
DB 232 LKTVITLVGVFIICMLPASITLID-STCPVRACPVLYKAHFAPATLNSLNPVITYW 290
OY 310 RSREVCRAVLSEFLCC---GCLRGLMKG--PG 335
DB 291 RSRDLAREVLRPLLCWRCKGKATGRGNGMPG 321

RESULT 6
EDG2 MOUSE STANDARD: PRT; 364 AA.
ID EDG2 MOUSE PRT; 364 AA.
AC 061130; P70420; P56487;
DT 01-NOV-1997 (Rel. 35; Created)
DT 15-JUL-1998 (Rel. 36; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Lysophosphatidic acid receptor (EDG-2) (RECL.3) (VZG-1).
GN EDG2 OR GPCR26 OR VZG1.
OS Mus musculus (Mouse) and
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090, 10116;
RN 11
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse; STRAIN=BALB/c;
RA MEDLINE=97081105; PubMed=8922387;
RA Hecht J.H., Weiner J.A., Post S.R., Chun J.;
RT "Ventricular zone gene-1 (vzg-1) encodes a lysophosphatidic acid
RT receptor expressed in neurogenic regions of the developing cerebral
RT cortex";
RL J. Cell Biol. 135:1071-1083(1996).
RN 12
RP SEQUENCE OF 19-364 FROM N.A.
RC SPECIES=Mouse; STRAIN=C57BL/6; TISSUE=Brain;
RA MEDLINE=97165887; PubMed=9013780;
RA Macre A.D., Premont R.T., Jaber M., Petersen A.S., Leftkowitz R.J.;
RT "Cloning, characterization, and chromosomal localization of recl.3, a
RT member of the G-protein-coupled receptor family highly expressed in
RT brain.";
RL Brain Res. Mol. Brain Res. 42:245-254(1996).
RN 13

```

```

RP SEQUENCE FROM N.A.
RC SPECIES=Rat; STRAIN=Sprague-Dawley; TISSUE=Olfactory bulb;
RA MEDLINE=98424094; PubMed=9753172;
RA Allard J., Barron S., Diaz J., Lubetzki C., Zalc B., Schwartz J.-C.,
RA Sokoloff P.;
RT "A rat G protein-coupled receptor selectively expressed in myelin-
RT forming cells.";
RL Eur. J. Neurosci. 10:1045-1053(1998).
CC -1- FUNCTION: RECEPTOR FOR LYOPHOSPHATIDIC ACID (LPA), A MEDIATOR OF
CC DIVERSE CELLULAR ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
CC EMBL: U70622; AAC52923.1; -.
CC EMBL: U48235; AAC53035.1; -.
CC EMBL: AF014418; AAB86381.1; -.
CC MGD: MGI:108429; Edg2.
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC InterPro: IPR004065; LPAREceptor.
CC Pfam: PF00001; 7tm_1; 1.
CC PRINTS: PR01527; LPARECEPTOR.
CC PROSITE: PS00237; G-PROTEIN_RECPT_FL1; 1.
CC PROSITE: PS0262; G-PROTEIN_RECPT_FL2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Phosphorylation; Lipoprotein; Palmitate.
CC K W
FT DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 51 75 1 (POTENTIAL).
FT DOMAIN 76 82 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 83 111 2 (POTENTIAL).
FT DOMAIN 112 125 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 126 144 3 (POTENTIAL).
FT DOMAIN 145 163 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 164 189 4 (POTENTIAL).
FT DOMAIN 190 205 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 206 226 5 (POTENTIAL).
FT DOMAIN 227 258 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 259 294 6 (POTENTIAL).
FT DOMAIN 281 294 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 295 315 7 (POTENTIAL).
FT DOMAIN 316 364 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 327 327 PALMITATE (BY SIMILARITY).
FT LIPID 327 327 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 35 35 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 225 225 Y -> S (IN REF. 2).
SQ SEQUENCE 364 AA; 41119 MW; B0FA6265A68887 CRC64;

Query Match 22.6%; Score 444.5; DB 1; Length 364;
Best Local Similarity 34.3%; Pred. No. 8,6e-23;
Matches 106; Conservative 51; Mismatches 141; Indels 11; Gaps 4;

OY 21 HSRLLVLYHNSGRLAGRGEPEDGGLALRLSVASCLVLENLVLVAATISMRRRW 80
DB 26 YNESIAFFYNRSGKYLATEWNTVSKL--VMGLITVCVFIMLLVLAAYVNRHFP 83
OY 81 VYCCVNTITSLDITGAAYLANVLSGARFTRLPAPQWFLREGILFTALASTSLFTFA 140
DB 84 IYILMANLAADFRAGLAFYFLMFTGNTRLRLVSTMLPGLIDISLRSASVANLLAIA 143
OY 141 GERFATVPRVAESGATRTSRVYFGICMLLALLGLMPLLGNCICAEDRCSSLLPLY 200
DB 144 IERHITVFRMQLHT-RMSNRRVVVYIYIMTAIYMGAIIPSGWNCICDIDHCSNMAPLY 202
OY 201 SKRYTLFCLVTFAGVATIMGLYALIFRVOASGOKAPRPAARRKAR-----LLKTVIM 255
DB 203 SDSYLVFAIFENLVTFVVAVVLYAHIFGYVRQRTMRSSRSHSGPRRNDTMMSLKTVI 262

```



```
DR EMBL: U80811; AAC51139.1; -
DR EMBL: Y09479; CAA70686.1; -
DR EMBL: Y09479; CAA70687.1; -
DR EMBL: U78192; AAC00530.1; -
DR Genew: HGNC:3166; EDG2.
DR MIM: 602282; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR004065; LPareceptor.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR01527; LPARECEPTOR.
DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 51 75 1 (POTENTIAL).
FT DOMAIN 76 82 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 83 111 2 (POTENTIAL).
FT DOMAIN 112 125 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 126 144 3 (POTENTIAL).
FT DOMAIN 145 163 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 164 189 4 (POTENTIAL).
FT DOMAIN 190 205 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 206 226 5 (POTENTIAL).
FT DOMAIN 227 258 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 259 280 6 (POTENTIAL).
FT DOMAIN 281 294 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 295 315 7 (POTENTIAL).
FT DOMAIN 316 364 CYTOPLASMIC (POTENTIAL).
FT LIPID 327 327 PALMITATE (BY SIMILARITY).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 340 340 S -> G (IN REF. 2).
SQ SEQUENCE 364 AA; 41139 MW; 59E7722FD00DE74 CRC64;
```

```
Query Match
Best Local Similarity 22.3%; Score 438; DB 1; Length 364;
Matches 111; Conservative 54; Mismatches 152; Indels 18; Gaps 6;
```

```
OY 1 MNATGPVAPESCOOLAAG-----HSRLIVYHNHSGR-LAGRGPDGGIGALRGISV 54
DB 1 MAISTSIIVISQPGFTMANEPQCFYNESIAFYNRSGKHLATEWNTVS---KLVMGIGI 57
OY 55 AASCLVIVENLLVLAITSHMSRRVYVCVNTITLSDLTGAAVLAVNLGARTFPLA 114
DB 58 TVCIFIMLANLLVVAIYVNRFRHPIYYLMANLAADFFGALVYFLMFTMGPTRLTJ 117
OY 115 PAQWFLREGILFTALASTFSILFTAGERPAMVPRVASEGATSKRYGFIGLWLLAA 174
DB 118 VSTWLLRQGLIDTSLTASVAVNLATAIERHITVFERMQLHT-RMSNRVVVIVVITWMAI 176
OY 175 LIGMLPLIGMNCICAFDRCSLLPLYSKRYILFCVIVPAGVLTIMGLYGAIFRLVOASG 234
DB 177 VMGAIPSYGMNCICDIENCSNMNAPLYSDSYLVFMAIFLVFVVMVVAIHAIFGVKORT 236
OY 235 OKAPPAARRRRR-----LKTIVIMILAFVCGPLGLLLADVFGSNIMAGEYLKGM 289
DB 237 MKMSHSGGPRRRRDTMMSLTKTIVYVIGAFIICWTBGLVLLDVCPCQCDVLAVER-- 294
OY 290 DWIILAVLNSAVNPDIYSFRSRECVGLSFLCC 324
DB 295 -FELLIAEFNSAMNPDIYSYDRKESATFRQILCC 328
```

```
RESULT 9
EDG2_BOVIN STANDARD; PRT; 364 AA.
AC Q28031.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Lysophosphatidic acid receptor (EDG-2) (Ref. 3).
GN EDG2.
```

```
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97165887; PubMed=9013780;
RA Macrae A.D., Premont R.T., Jaber M., Petersen A.S., Lefkowitz R.J.,
RT "Cloning, characterization, and chromosomal localization of rec1.3, a
RT member of the G-protein-coupled receptor family highly expressed in
RT brain.";
RL Brain Res. Mol. Brain Res. 42:245-254(1996).
CC -!- FUNCTION: RECEPTOR FOR LYSOPHOSPHATIDIC ACID (LPA), A MEDIATOR OF
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL: U48236; AAC4695.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR004065; LPareceptor.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR01527; LPARECEPTOR.
DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 51 75 1 (POTENTIAL).
FT DOMAIN 76 82 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 83 111 2 (POTENTIAL).
FT DOMAIN 112 125 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 126 144 3 (POTENTIAL).
FT DOMAIN 145 163 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 164 189 4 (POTENTIAL).
FT DOMAIN 190 205 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 206 226 5 (POTENTIAL).
FT DOMAIN 227 258 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 259 280 6 (POTENTIAL).
FT DOMAIN 281 294 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 295 315 7 (POTENTIAL).
FT DOMAIN 316 364 CYTOPLASMIC (POTENTIAL).
FT LIPID 327 327 PALMITATE (BY SIMILARITY).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 364 AA; 41070 MW; DE94675BA407B2F6 CRC64;
```

```
Query Match
Best Local Similarity 22.0%; Score 432.5; DB 1; Length 364;
Matches 103; Conservative 51; Mismatches 140; Indels 11; Gaps 4;
```

```
OY 25 IVLHYNHSGRLAAGRGPDGGIGALRGLSAASCLVLENLVLVAITSHMSRRVYVC 84
DB 30 IAFVYNRSGKYLATENWNTYKL--VNGLGITVCIIFIMLANLLVMAIYVNRFRHPIYLL 87
OY 85 LVNITISDLTGAAVLAVNLGSGARTFRLAPQWFLRBSLFTALAASTFSILFTAGERE 144
DB 88 MANLIAADFFAGLAFYLYMFNTPNTRRLTVSWLLRQGLIDTSLVSVANLATAIERH 147
OY 145 ATWVRVPAESGATKTSRYVGFIGLWLLAALGMLPLIGMNCICAFDRCSLLPLYSKRY 204
DB 148 IITFR-MQHLARNSNRVVVIVITWMAIYMGALPSYGMNCICDIENCSNMNAPLYSDSY 206
OY 205 IIFCLVIVPAGVLTIMGLYGAIFRLVOASGOKAPPAARRRRR-----LKTIVIMILAA 259
```



```

DB 207 LVEFMAIFNLVTFVVVAVVIAHIFGVYVORTMRMSSHGSPRRNDPTMMSLTKTVIYVIGA 266
OY 260 FLVCMGPFLLGLLADVFGSNLMAOETLKGMDMILALAVLSAVNPITYSPRSREVCRAVL 319
DB 267 FLICHTPGVILLVLLVCCPGCDVLAVER---FFLLAFENSMNPITYSYDKEMSAFER 323
OY 320 SFLLCC 324
DB 324 QILCC 328

RESULT 10
CBIR_MOUSE STANDARD: PRT: 473 AA.
ID CBIR_MOUSE
AC P47746;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cannabinoid receptor 1 (CBI) (CB-R) (Brain-type cannabinoid receptor).
GN CNRL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=96272305; PubMed=8777318;
RA Chakrabarti A., Onalvi E.S., Chaudhuri G.;
RT "Cloning and sequencing of a cDNA encoding the mouse brain-type
RT cannabinoid receptor protein."
RL DNA Seq. 3:385-388(1995).
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=129;
RA Bonner T.I.;
RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
RN [3]
RC SEQUENCE FROM N.A.
RX MEDLINE=96429553; PubMed=8832654;
RA Ho B.-Y., Zhao J.;
RT "Determination of the cannabinoid receptors in mouse x rat hybridoma
RT NG108-15 cells and rat GH4C1 cells."
RL Neurosci. Lett. 212:123-126(1996).
RN [4]
RC SEQUENCE FROM N.A.
RA Yuan Z.-Q., Li L., Qiu B.-S., Song D.-K.;
RT "cDNA cloning and expression analysis of mouse cannabinoid receptor
RT (CBI) gene."
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
RN [5]
RC SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=99105999; PubMed=9888857;
RA Ledert C., Valverde O., Cosu G., Petitot F., Aubert J.F., Beslot F.,
RA Boehme G.A., Imperato A., Pedrazzini T., Roques B.P., Vassart G.,
RA Fratta W., Parmentier M.;
RT "Unresponsiveness to cannabinoids and reduced addictive effects of
RT opiates in CBI receptor knockout mice."
RL Science 283:401-404(1999).
CC -1- FUNCTION: INVOLVED IN CANNABINOID-INDUCED CNS EFFECTS. ACTS BY
CC INHIBITING ADENYLATE CYCLASE. COULD BE A RECEPTOR FOR ANANDAMIDE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC -----
DR EMBL: U17985; AAA57202.1; -
DR EMBL: U22948; AAA54413.1; -
DR EMBL: U40709; AAA91176.1; -
DR EMBL: AF153345; AAD34624.1; -
DR EMBL: Y18374; CAB42647.1; -
DR MGD: MGI:104615; Cnrl.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 117
FT TRANSSEM 118 143
FT DOMAIN 144 155
FT TRANSSEM 156 176
FT DOMAIN 177 188
FT TRANSSEM 189 213
FT DOMAIN 214 233
FT TRANSSEM 234 256
FT DOMAIN 257 274
FT TRANSSEM 275 300
FT DOMAIN 301 345
FT TRANSSEM 346 366
FT DOMAIN 367 378
FT TRANSSEM 379 400
FT DOMAIN 401 473
FT CARBOHYD 78 78
FT CARBOHYD 84 84
FT CONFLICT 9 9
FT CONFLICT 115 115
FT CONFLICT 211 211
SQ SEQUENCE 473 AA; 52830 MW; E504168191CB6429 CRC64;
Query Match 17.98; Score 351; DB 1; Length 473;
Best Local Similarity 30.18; Pred. No. 1,6e-16;
Matches 105; Conservative 61; Mismatches 129; Indels 54; Gaps 11;
```

```

OY 48 ALRGISVAASCLVVENLVLAATITSHMRSRR--VYCVLVNITLSDLTGAAYLVANVL 105
DB 119 ALAVSLGLGFTVLENLVLCVIL-HSRILRCRPSHFISGLAVADLGSVIFYVSFD 177
OY 106 SGARTFRLPAQWFLREGLFTLALASTFSLFTAGSRFAVMPVPA-ESGATKTSVYG 164
DB 178 FHFVFRKDSPVVFLFKLGVTASFTASVGSFLTAIDRYISIHRLAYKRIIVRPKAVA 237
OY 165 FTGLCMILALLGMPLILGNMCLCAPDRCSLLPYLSKRYLPCTVFAGLVATIMLYG 224
DB 238 FC-LMWTIAIVIAVPLGMCKKLSQVSDIFPLIDETVLMFVIGTVSYLLFIYAYM 296
OY 225 AIF-----RLVQASGOKA-----PRPAARRKARLLKTVLMLIAFLVC 263
DB 297 YILMKASHAVRMIGRGQKSIITHSBDKVVQYTRPDQAMODIRLAKTVLLIVLVIIIC 356
OY 264 WGPLFGLLADVFGSNLMAOETLKGMDMIL-----ALAVLSAVNPITYSPRSREVC 315
DB 357 WGPLLAIVYDVFGK-----MNLKITYPAFCSMCLLNSTVNPITYALRSKDLR 406
OY 316 RAVLSFL--CCGCLRL--GMRPGDCLARAFAVHSGASTDTSIRPDS 360
DB 407 HAFNSMPSCSGTQAPLPDLSMGDSDDL-----HKHANNTASMRRAES 449

RESULT 11
CBIR_RAT STANDARD: PRT: 473 AA.
ID CBIR_RAT
AC P20272;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cannabinoid receptor 1 (CBI) (CB-R) (Brain-type cannabinoid receptor).
GN CNRL OR SKR6.
OS Rattus norvegicus (Rat).
```



```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cortex;
RX MEDLINE=9033039; PubMed=2165569;
RA Matsuda L.A., Lolait S.J., Brownstein M.J., Young A.C., Bonner T.I.;
RT "Structure of a cannabinoid receptor and functional expression of the
RT cloned cDNA.";
RL Nature 346:561-564(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96429553; PubMed=8832654;
RA Ho B.Y., Zhao J.;
RT "Determination of the cannabinoid receptors in mouse x rat hybridoma
RT NG108-15 cells and rat GH41 cells.";
RL Neurosci. Lett. 212:123-126(1996).
CC -1- FUNCTION: INVOLVED IN CANNABINOID-INDUCED CNS EFFECTS. ACTS BY
CC -1- INHIBITING ADENYLATE CYCLASE. COULD BE A RECEPTOR FOR ANANDAMIDE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; X55812; CAA9332.1; -
CC DR EMBL; U40395; AAA99067.1; -
CC DR PIR; A33117; A33117.
CC DR InterPro: IPR000276; GPCR_Rhodpsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein.
CC FT DOMAIN 1 117
CC FT TRANSSEM 118 143
CC FT DOMAIN 144 155
CC FT TRANSSEM 156 176
CC FT DOMAIN 177 188
CC FT TRANSSEM 189 213
CC FT DOMAIN 214 233
CC FT TRANSSEM 234 256
CC FT DOMAIN 257 274
CC FT TRANSSEM 275 300
CC FT DOMAIN 301 345
CC FT TRANSSEM 346 366
CC FT DOMAIN 367 378
CC FT TRANSSEM 379 400
CC FT DOMAIN 401 473
CC FT CARBOHYD 78
CC FT CARBOHYD 84
CC SEQUENCE 473 AA; 52845 MW; E59A66AFEL7B646C CRC64;
Query Match 17.9%; Score 351; DB 1; Length 473;
Best Local Similarity 30.1%; Pred. No. 1.6e-16;
Matches 105; Conservative 61; Mismatches 129; Indels 54; Gaps 11;
```

```
OY 225 AIF-----RLVQASGQRA-----PPAARARRRLKTVLITLAFVLC 263
DB 297 YILMKASHAVRMIORGTOXSIITHTSEDKVOYTRPDQARMDIRAKTVLITLVLLIC 356
OY 264 WGPLEGLLADVFGSNMAQGYLRGMDIL-----ALAVNSAVNPDIYSFRSREVC 315
DB 337 WGPLAIIVYVYFGK-----MKLIKTYAFQSMCLLNSVNPFIYALFRSKDR 406
OY 316 RAVLSFL--CCGCLRL--GMRPGDCIARAVEAHSGASTDSSLRPDS 360
DB 407 HAFRSMFPCSGEATQPLDNSWGDSDCL-----HKHANNATSMHRAAES 449
RESULT 12
CB1B_FUGRU STANDARD; PRT; 470 AA.
AC 098895;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cannabinoid receptor type 1B.
GN CB1B.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=97001167; PubMed=8812500;
RA Yamaguchi F., Macrae A., Brenner S.;
RT "Molecular cloning of two cannabinoid type 1-like receptor genes from
RT the puffer fish Fugu rubripes.";
RL Genomics 35:603-605(1996).
CC -1- FUNCTION: INVOLVED IN CANNABINOID-INDUCED CNS EFFECTS (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; X94402; CAA64175.1; -
CC DR InterPro: IPR000276; GPCR_Rhodpsn.
CC DR Pfam: PF00001; 7tm_1; 1.
CC DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein.
CC FT DOMAIN 1 113
CC FT TRANSSEM 114 139
CC FT DOMAIN 140 151
CC FT TRANSSEM 152 172
CC FT DOMAIN 173 184
CC FT TRANSSEM 185 209
CC FT DOMAIN 210 229
CC FT TRANSSEM 230 252
CC FT DOMAIN 253 270
CC FT TRANSSEM 271 296
CC FT DOMAIN 297 341
CC FT TRANSSEM 342 362
CC FT DOMAIN 363 374
CC FT TRANSSEM 375 396
CC FT DOMAIN 397 470
CC FT CARBOHYD 78
CC FT CARBOHYD 86
CC SEQUENCE 470 AA; 52081 MW; CE87CD37FDF9192 CRC64;
```

|                       |  |              |            |                                     |
|-----------------------|--|--------------|------------|-------------------------------------|
| FT                    | DOMAIN   | 176          | 187        | EXTRACELLULAR (POTENTIAL).          |
| FT                    | TRANSMEM   | 188          | 212        | 3 (POTENTIAL).                      |
| FT                    | DOMAIN   | 213          | 232        | CYTOPLASMIC (POTENTIAL).            |
| FT                    | TRANSMEM   | 233          | 253        | 4 (POTENTIAL).                      |
| FT                    | DOMAIN   | 256          | 273        | EXTRACELLULAR (POTENTIAL).          |
| FT                    | TRANSMEM   | 274          | 299        | 5 (POTENTIAL).                      |
| FT                    | DOMAIN   | 300          | 344        | CYTOPLASMIC (POTENTIAL).            |
| FT                    | TRANSMEM   | 345          | 365        | 6 (POTENTIAL).                      |
| FT                    | DOMAIN   | 366          | 377        | EXTRACELLULAR (POTENTIAL).          |
| FT                    | TRANSMEM   | 378          | 399        | 7 (POTENTIAL).                      |
| FT                    | DOMAIN   | 400          | 472        | CYTOPLASMIC (POTENTIAL).            |
| FT                    | CARBOHYD   | 77           | 77         | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT                    | CARBOHYD   | 83           | 83         | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| SO                    | SEQUENCE   | 472 AA;      | 52887 MW;  | A2774DBF8F2DAF34 CRC64;             |
| Query Match           |  |              |            |                                     |
| Best Local Similarity |  | 17.5%;       | Score 345; | DB 1; Length 472;                   |
| Matches 102;          |  | Conservative | 60;        | Mismatches 124; Indels 54; Gaps 11. |
| QY                    | 48 ALRGLSVAACGLVLENLVLAITSHMSRKM--VYICLVNTILSDLLGAIYLANVLL                 | 105          |            |                                     |
| DB                    | 118 AIAVLSLTIGTFVLENLVLCVIL-HSRSLRCRSPYHGISLAVADLGSVFEVYSFD                | 176          |            |                                     |
| QY                    | 106 SGATFFRIAAQWFLEREGLLFTLAASTSLFTTGEFFAPVAPVA-ESGATKTSRYVG               | 164          |            |                                     |
| DB                    | 177 FHVHRKDSQNVNLFPIKIGVTASFTASVGSLETLIDKRIYSIHRRPLAKYIVRPAAVYA            | 236          |            |                                     |
| QY                    | 165 FTIGCMLLAALLGMPLPLGNCLEAFDRSSILPLYSKRYILFCLVIFPAQVLAITMGLYG            | 224          |            |                                     |
| DB                    | 237 FC-LMNTIAIVAVLPLLLGMCNKLOSVDIEFILDITVLMFNIQVTSVLLFEIVAYM               | 295          |            |                                     |
| QY                    | 225 AIF-----RLVQASGOKA-----PRPARARRARLKTVMILLAFLVC                         | 263          |            |                                     |
| DB                    | 296 YILKHAHHAVRMIGTQKSIITHSEDGQVYRRPQAMDIRLAKTLVLLVLLIIC                   | 355          |            |                                     |
| QY                    | 264 WGPFLGLLADVGSNLMAQETIRGMDWL-----ALAVLNSAVNPLTIFSRREVC                  | 315          |            |                                     |
| DB                    | 356 WGPLLAIVMVDVEGK-----MNKLIKTFEFCMCLLNSVNPITVALRSKDLR                    | 405          |            |                                     |
| QY                    | 316 RAVLSFL--CCGCLRL--GMRGPGDCLARAEAHSGASTY                                | 351          |            |                                     |
| DB                    | 406 HAFRSMFPSCGTAQPLDMSMGDSPDL-----HKHANNT                                 | 439          |            |                                     |
| RESULT 14             |  |              |            |                                     |
| EDGL_MOUSE            |  |              |            |                                     |
| AC                    | EDGL_MOUSE   | STANDARD:    | PRT:       | 180 AA.                             |
| AC                    | P52592;  |              |            |                                     |
| DT                    | 01-OCT-1996 (Rel. 34, Created)   |              |            |                                     |
| DT                    | 01-OCT-1996 (Rel. 34, Last sequence update)                                |              |            |                                     |
| DT                    | 15-JUL-1998 (Rel. 36, Last annotation update)                              |              |            |                                     |
| DE                    | Probable G protein-coupled receptor EDG-1 like (Fragment).                 |              |            |                                     |
| GN                    | GPR13.   |              |            |                                     |
| OS                    | Mus musculus (Mouse).  |              |            |                                     |
| OC                    | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;          |              |            |                                     |
| OC                    | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.         |              |            |                                     |
| OX                    | NCBI_Taxid:10090;  |              |            |                                     |
| OX                    | [1]  |              |            |                                     |
| RP                    | SEQUENCE FROM N.A.   |              |            |                                     |
| RC                    | TISSUE=Testis;   |              |            |                                     |
| RC                    | MEDLINE=94116980; PubMed=8288218;  |              |            |                                     |
| RA                    | Wilkie T.M., Chen Y., Gilbert D.J., Moore K.J., Yu L., Simon M.I.,         |              |            |                                     |
| RA                    | Copeland N.G., Jenkins N.A.;   |              |            |                                     |
| RT                    | "Identification, chromosomal   |              |            |                                     |
| RT                    | mammalian G-protein-coupled receptors";                                    |              |            |                                     |
| RL                    | Genomics 18:175-184(1993).   |              |            |                                     |
| CC                    | -1- SUBCELLULAR LOCATION: Integral membrane protein.                       |              |            |                                     |
| CC                    | -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLD RECEPTORS.         |              |            |                                     |
| CC                    | This SWISS-PROT entry is copyright. It is produced through a collaboration |              |            |                                     |
| CC                    | between the Swiss Institute of Bioinformatics and the EMBL outstation -    |              |            |                                     |
| CC                    | the European Bioinformatics Institute. There are no restrictions on its    |              |            |                                     |
| CC                    | use by non-profit institutions as long as its content is in no way         |              |            |                                     |



|    |          |                   |  |
|----|----------|-------------------|--|
| FT |          |                   | GFEPKQFPLTSRSGSPQEKMTAGDNPOLPADQVNTTEI |
| FT |          |                   | NKSISSF -> MALQIPPSAPSLTSCCTMAQMTSTYS  |
| FT |          |                   | (1IN SHORT ISOFORM).                   |
| FT | CONFLICT | 200               | F -> L (1IN REF. 4).                   |
| FT | CONFLICT | 216               | I -> V (1IN REF. 4).                   |
| FT | CONFLICT | 246               | V -> A (1IN REF. 4).                   |
| EQ | SEQUENCE | 472 AA: 52857 MW: | 1DSE49061D12ABF2 CRC64;                |



|                           |   |
|---------------------------|---|
| FT                        | /note= "transmembrane domain IV"                                    |
| FT                        | 183..199  |
| FT                        | /note= "extracellular domain EL-II"                                 |
| FT                        | 200..224  |
| FT                        | /note= "transmembrane domain V"                                     |
| FT                        | 225..250  |
| FT                        | /note= "intracellular domain IL-III"                                |
| FT                        | 251..272  |
| FT                        | /note= "transmembrane domain VI"                                    |
| FT                        | 273..285  |
| FT                        | /note= "extracellular domain EL-II"                                 |
| FT                        | 286..304  |
| FT                        | /note= "transmembrane domain VII"                                   |
| FT                        | 2   |
| FT                        | /note= "N-glycosylated"   |
| FT                        | 30  |
| FT                        | /note= "N-glycosylated"   |
| FT                        | 77  |
| FT                        | /note= "O-phosphorylated"   |
| FT                        | 159   |
| FT                        | /note= "O-phosphorylated"   |
| FT                        | 308   |
| FT                        | /note= "O-phosphorylated"   |
| FT                        | 360   |
| FT                        | /note= "O-phosphorylated"   |
| FT                        | 380   |
| FT                        | /note= "O-phosphorylated"   |
| PT                        |   |
| ET                        |   |
| XX                        |   |
| PB                        | WO935106-A2.  |
| XX                        |   |
| PD                        | 15-JUL-1999.  |
| XX                        |   |
| XX                        |   |
| PF                        | 30-DEC-1998; 98WO-CA01196.  |
| XX                        |   |
| PR                        | 30-DEC-1997; 97US-0070184.  |
| XX                        |   |
| PA                        | (ALIX ) ALLELIX BIOPHARMACEUTICALS INC.                             |
| XX                        |   |
| P1                        | Gupta AK, Munroe DG, Zastawny RL;                                   |
| DR                        | WPI: 1999-419322/35.  |
| DR                        | N-PDB: AAX59366.  |
| XX                        |   |
| PT                        | A nucleic acid sequence encoding human EDG-7 receptor, useful for   |
| XX                        | treating disorders associated with aberrant expression              |
| PS                        | Claim 12; Fig 2A; 72pp; English.                                    |
| XX                        |   |
| CC                        | The present sequence represents a novel human homologue of the      |
| CC                        | EDG-7 receptor, i.e. HEDG-7, a 7-transmembrane G protein coupled    |
| CC                        | receptor. The sequence is predicted from isolated BAC and PAC       |
| CC                        | clones (see AAX59366) and differs at only 2 amino acid residues     |
| CC                        | (140 and 178) from the sequence (see AAU06412) from an isolated     |
| CC                        | cDNA clone. Also claimed are an expression vector comprising a      |
| CC                        | HEDG-7 nucleotide sequence, a host cell, and methods of using       |
| CC                        | HEDG-7 receptor to identify potential HEDG-7 ligands and            |
| CC                        | antagonists. Ligands, agonists and antagonists of HEDG-7 may        |
| CC                        | be useful for detecting inflammation or disease associated with     |
| CC                        | abnormal levels of HEDG-7 expression. Detection of aberrant         |
| CC                        | expression of HEDG-7 can accelerate diagnosis and proper treatment  |
| CC                        | of abnormal conditions, e.g. adult respiratory distress, asthma,    |
| CC                        | rheumatoid arthritis, cardiac ischaemia, acute pancreatitis, septic |
| CC                        | shock, psoriasis, acute cyclosporin nephrotoxicity and early        |
| CC                        | diabetic glomerulopathy, as well as lung damage following exposure  |
| CC                        | to cigarette smoke, asbestos or silica.                             |
| XX                        |   |
| SQ                        | Sequence 384 AA;  |
| Query Match               | 100.0%; Score 1966; DB 20; Length 384;                              |
| Best Local Similarity     | 100.0%; Pred. No. 2.4e-206;   |
| Matches 384; Conservative | 0; Mismatches 0; Indels 0; Gaps                                     |

|          |  |   |     |
|----------|--|---|-----|
| Db       | 1  | MAATGTPAAPESCQDLAAGSHRLIVLHNSGRLAERGGEDGGALRGSLVAASCLV      | 60  |
| QY       | 61   | VLENLVLVAATISSHMRSRRWYYCLVNTTSDLTLTGAAYLANVLLSGARTFRLAPQWFL | 120 |
| Db       | 61   | VLENLVLVAATISSHMRSRRWYYCLVNTTSDLTLTGAAYLANVLLSGARTFRLAPQWFL | 120 |
| QY       | 121  | REGLLFTFLAASSTSLFTFAEGERATWVRPAEAGAKTTSRYGFIQLCMLAALLGMLP   | 180 |
| Db       | 121  | REGLLFTFLAASSTSLFTFAEGERATWVRPAEAGAKTTSRYGFIQLCMLAALLGMLP   | 180 |
| QY       | 181  | LLGWMCLAFDRCSSSLPEYKSKRIELCVIFAGVLATIMGILGATFRLVQASGQKAPP   | 240 |
| Db       | 181  | LLGWMCLAFDRCSSSLPEYKSKRIELCVIFAGVLATIMGILGATFRLVQASGQKAPP   | 240 |
| QY       | 241  | AARRARRLKTVLMILLALFLVCGPLFGLLADVFSGSNMAOEYLRGMDWITALAVLNS   | 300 |
| Db       | 241  | AARRARRLKTVLMILLALFLVCGPLFGLLADVFSGSNMAOEYLRGMDWITALAVLNS   | 300 |
| QY       | 301  | AVNPITTFRSREVCRAVLAFLCCGCLRLGRRGGDCLARVAHSGASTTSSLRPRS      | 360 |
| Db       | 301  | AVNPITTFRSREVCRAVLAFLCCGCLRLGRRGGDCLARVAHSGASTTSSLRPRS      | 360 |
| QY       | 361  | FRGSRSLFRMRREPLSSISVSRSI 384                                |     |
| Db       | 361  | FRGSRSLFRMRREPLSSISVSRSI 384                                |     |
| RESULT 2 |  |   |     |
| AA69500  |  |   |     |
| ID       | AA69500 standard; Protein; 384 AA.                                     |   |     |
| XX       |  |   |     |
| AC       | AA69500;   |   |     |
| XX       |  |   |     |
| DT       | 12-JUL-2000 (first entry)  |   |     |
| XX       |  |   |     |
| DE       | A human G-protein coupled receptor designated 14275 receptor.          |   |     |
| XX       |  |   |     |
| KW       | Human; G-protein coupled receptor; GPCR; EDG receptor; 14275 receptor; |   |     |
| KW       | signalling pathway; GPCR-mediated disorder; anaemia; neutropenia;      |   |     |
| KW       | thrombocytopenia; inflammation.  |   |     |
| XX       |  |   |     |
| OS       | Homo sapiens.  |   |     |
| XX       |  |   |     |
| FH       | Key  | Location/Qualifiers   |     |
| FT       | Domain   | 1..50   |     |
| FT       |  | /note= "extracellular domain"                               |     |
| FT       | Modified-site  | 2..5  |     |
| FT       |  | /note= "N glycosylation site"                               |     |
| FT       | Modified-site  | 30..33  |     |
| FT       |  | /note= "N glycosylation site"                               |     |
| FT       | Modified-site  | 32..34  |     |
| FT       |  | /note= "transmembrane segment"                              |     |
| FT       | Modified-site  | 44..49  |     |
| FT       |  | /note= "protein kinase C phosphorylation site"              |     |
| FT       | Modified-site  | 51..56  |     |
| FT       |  | /note= "N-myristoylation site"                              |     |
| FT       | Domain   | 51..331   |     |
| FT       |  | /note= "transmembrane domain"                               |     |
| FT       | Domain   | 51..71  |     |
| FT       |  | /note= "transmembrane segment"                              |     |
| FT       | Modified-site  | 77..79  |     |
| FT       |  | /note= "protein kinase C phosphorylation site"              |     |
| FT       | Region   | 81..105   |     |
| FT       |  | /note= "transmembrane segment"                              |     |
| FT       | Modified-site  | 87..90  |     |
| FT       |  | /note= "N glycosylation site"                               |     |
| FT       | Modified-site  | 89..92  |     |
| FT       |  | /note= "casein kinase II phosphorylation site"              |     |
| FT       | Modified-site  | 110..112  |     |
| FT       |  | /note= "protein kinase C phosphorylation site"              |     |
| FT       | Region   | 123..141  |     |

|    |   |   |
|----|---|---|
| FT | Modified-site   | /note="transmembrane segment"                 |
| FT | Modified-site   | 123..128                                      |
| FT | Modified-site   | /note="N-myristoylation site"                 |
| FT | Modified-site   | 139..142                                      |
| FT | Region  | /note="casein kinase II phosphorylation site" |
| FT | Modified-site   | 142..144                                      |
| FT | Modified-site   | /note="GPCR signal transduction signature"    |
| FT | Modified-site   | 155..160                                      |
| FT | Modified-site   | /note="N-myristoylation site"                 |
| FT | Modified-site   | 159..161                                      |
| FT | Region  | /note="protein kinase C phosphorylation site" |
| FT | Modified-site   | 162..184                                      |
| FT | Modified-site   | /note="transmembrane segment"                 |
| FT | Modified-site   | 201..203                                      |
| FT | Region  | /note="protein kinase C phosphorylation site" |
| FT | Modified-site   | 204..227                                      |
| FT | Modified-site   | /note="transmembrane segment"                 |
| FT | Modified-site   | 214..219                                      |
| FT | Modified-site   | /note="N-myristoylation site"                 |
| FT | Modified-site   | 221..226                                      |
| FT | Region  | /note="N-myristoylation site"                 |
| FT | Region  | 253..276                                      |
| FT | Modified-site   | /note="transmembrane segment"                 |
| FT | Modified-site   | 269..274                                      |
| FT | Region  | /note="N-myristoylation site"                 |
| FT | Modified-site   | 291..331                                      |
| FT | Modified-site   | /note="transmembrane segment"                 |
| FT | Modified-site   | 308..310                                      |
| FT | Modified-site   | /note="protein kinase C phosphorylation site" |
| FT | Modified-site   | 347..352                                      |
| FT | Modified-site   | /note="N-myristoylation site"                 |
| FT | Modified-site   | 349..352                                      |
| FT | Modified-site   | /note="casein kinase II phosphorylation site" |
| FT | Modified-site   | 354..356                                      |
| FT | Modified-site   | /note="protein kinase C phosphorylation site" |
| FT | Modified-site   | 360..362                                      |
| FT | Modified-site   | /note="protein kinase C phosphorylation site" |
| FT | Modified-site   | 368..370                                      |
| FT | Modified-site   | /note="protein kinase C phosphorylation site" |
| FT | Modified-site   | 380..382                                      |
| FT | Modified-site   | /note="protein kinase C phosphorylation site" |
| XX |   |   |
| PN | WO200014233-A1.   |   |
| PD | 16-MAR-2000.  |   |
| XX |   |   |
| PF | 03-SEP-1999;  | 99WO-US20347.                                 |
| XX |   |   |
| PR | 03-SEP-1998;  | 98US-0146416.                                 |
| PR | 03-SEP-1999;  | 99US-0390039.                                 |
| XX |   |   |
| PA | (MILL-) MILLENNIUM PHARM INC.   |   |
| XX |   |   |
| PI | Glucksmann MA, Hodge MR;  |   |
| XX | WPI: 2000-256983/22.  |   |
| DR | N-FSDB: AA299797.   |   |
| XX |   |   |
| PT | New G-protein coupled receptor used in receptor assays as a target for    |   |
| PT | diagnosis and treatment of receptor-mediated disorders including          |   |
| PT | anaemia, neutropenia or thrombocytopenia or a disorder involving          |   |
| PT | inflammation -  |   |
| XX |   |   |
| PS | Claim 1; Fig 1A-B; 117pp; English.  |   |
| XX |   |   |
| CC | The present sequence represents a human G-protein coupled receptor (GPCR) |   |
| CC | which is related to the EDG receptor family. The protein is designated    |   |
| CC | the 14275 receptor. The 14275 receptor participates in signalling         |   |
| CC | pathways. The 14275 polypeptide and polynucleotide are used as reagents   |   |
| CC | or targets in a receptor assay for treatment and diagnosis of             |   |
| CC | GPCR-mediated disorders. The disorders include anaemia, neutropenia or    |   |
| CC | thrombocytopenia or a disorder involving inflammation.                    |   |
| XX |   |   |

| Seq      | Sequence   | 384 AA:                    | 100.0%;       | Score 1966;         | DB 21;  | Length 384; |
|----------|--|----------------------------|---------------|---------------------|---------|-------------|
|          | Query Match  |                            | 100.0%;       | Pred. No. 2.4e-206; |         |             |
|          | Best Local Similarity  |                            | 100.0%;       | Pred. No. 2.4e-206; |         |             |
|          | Matches 384;   | Conservative 0;            | Mismatches 0; | Indels 0;           | Gaps 0; |             |
| QY       | 1 MNATGTPVAPESCOOLAAAGHSRLITVLVHNNHSGRLAGRGPEGGGLGALRGLSVASACTIV       | 60                         |               |                     |         |             |
| Db       | 1 MNATGTPVAPESCOOLAAAGHSRLITVLVHNNHSGRLAGRGPEGGGLGALRGLSVASACTIV       | 60                         |               |                     |         |             |
| QY       | 61 VLENLTVLAATSHMSRRWVYCCIVNTITLSDLLTGAAYLANVLLSGARTFRLPAQWFL          | 120                        |               |                     |         |             |
| Db       | 61 VLENLTVLAATSHMSRRWVYCCIVNTITLSDLLTGAAYLANVLLSGARTFRLPAQWFL          | 120                        |               |                     |         |             |
| QY       | 121 REGLLFTALAASTFSLITFAGERFATWVRPAVESGATKTSVYVFGICLCLWLAALLGMP        | 180                        |               |                     |         |             |
| Db       | 121 REGLLFTALAASTFSLITFAGERFATWVRPAVESGATKTSVYVFGICLCLWLAALLGMP        | 180                        |               |                     |         |             |
| QY       | 181 LLGNCCAPDCRSSLLPIYSKRYITFCVIFPAGVLAITMLYGAIFRLVQAAGQAKRP           | 240                        |               |                     |         |             |
| Db       | 181 LLGNCCAPDCRSSLLPIYSKRYITFCVIFPAGVLAITMLYGAIFRLVQAAGQAKRP           | 240                        |               |                     |         |             |
| QY       | 241 AARRKARLLKTYVLMILLAEVYCWGPFLFGLLIADYFGSINMAQEYLRGMDMILLAVLNS       | 300                        |               |                     |         |             |
| Db       | 241 AARRKARLLKTYVLMILLAEVYCWGPFLFGLLIADYFGSINMAQEYLRGMDMILLAVLNS       | 300                        |               |                     |         |             |
| QY       | 301 AVNPIIYSFSSREVCRAVLSFLCCGCLRIGMPGPDCLARAVEHSGASTTDSLRPRS           | 360                        |               |                     |         |             |
| Db       | 301 AVNPIIYSFSSREVCRAVLSFLCCGCLRIGMPGPDCLARAVEHSGASTTDSLRPRS           | 360                        |               |                     |         |             |
| QY       | 361 FRGSRSLSPMRREPLSSISSVRSI   | 384                        |               |                     |         |             |
| Db       | 361 FRGSRSLSPMRREPLSSISSVRSI   | 384                        |               |                     |         |             |
| RESULT 3 |  |                            |               |                     |         |             |
| ID       | AAW90862   | standard; Protein; 384 AA. |               |                     |         |             |
| AC       | AAW90862:  |                            |               |                     |         |             |
| XX       | 29-JUN-2000 (first entry)  |                            |               |                     |         |             |
| DE       | Human edg6 protein.  |                            |               |                     |         |             |
| KM       | edg6; human; G-coupled receptor; endothelial differentiation gene;     |                            |               |                     |         |             |
| KW       | antiinflammatory; immunomodulatory; antimicrobial; antiallergic;       |                            |               |                     |         |             |
| KM       | cytostatic; gene therapy; inflammation; autoimmune disease; allergy;   |                            |               |                     |         |             |
| OS       | tumour; Leukemia; lymphoma.  |                            |               |                     |         |             |
| OS       | Homo sapiens.  |                            |               |                     |         |             |
| PN       | DE19846979-A1.   |                            |               |                     |         |             |
| PD       | 23-MAR-2000.   |                            |               |                     |         |             |
| PF       | 13-OCT-1998; 98DE-1046979.   |                            |               |                     |         |             |
| PR       | 11-SEP-1998; 98DE-1043240.   |                            |               |                     |         |             |
| PA       | (DELB-) DELBRUECK CENT MOLECULARE MEDIZIN MAX.                         |                            |               |                     |         |             |
| FI       | Graepler M, Bernhardt G, Lipp M;                                       |                            |               |                     |         |             |
| DR       | WPI; 2000-258069/23.   |                            |               |                     |         |             |
| DR       | N-PSDB; AAX82789.  |                            |               |                     |         |             |
| PT       | New human and murine G-coupled receptor EDG (endothelial               |                            |               |                     |         |             |
| PT       | differentiation gene) 6, useful for modulating inflammatory and immune |                            |               |                     |         |             |
| PT       | reactions and for treatment of allergy or tumors                       |                            |               |                     |         |             |
| PS       | Claim 1; Page 6; 12pp; German.   |                            |               |                     |         |             |

CC This invention describes novel human and murine G-coupled receptors EDG  
CC (endothelial differentiation gene) 6 (I and II). The products of the  
CC invention have antiinflammatory, immunomodulatory, antimicrobial,  
CC antiallergic and cytostatic activity. (I) and (II) are involved in signal  
CC transduction. (I), (II) and their fragments, variants and mutants or  
CC binding partners, are used therapeutically to modulate the function of  
CC blood and body cells, particularly for inhibition of acute and chronic  
CC inflammation and to raise specific antibodies against them. They are used  
CC as a source of diagnostic oligonucleotides and for gene therapy.  
CC Antibodies against (I) and (II) are useful for diagnosis and optionally  
CC when coupled to therapeutic agents, toxins or other antibodies, to  
CC modulate immune and inflammatory responses for example immunological  
CC defects such as inflammation, infection, autoimmune diseases, allergy,  
CC tumors, leukemia and lymphoma. This sequence represents the human EDG6  
CC protein described in the method of the invention.

CC  
SQ Sequence 384 AA:

Query Match 100.0%; Score 1966; DB 21; Length 384;  
Best Local Similarity 100.0%; Pred. No. 2.4e-206;  
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNATGTPVAPESCOQLAAGHSRLIVLHYNHSGRLAGRGPEDEGLALRGLSYAASCLV 60  
DB 1 MNATGTPVAPESCOQLAAGHSRLIVLHYNHSGRLAGRGPEDEGLALRGLSYAASCLV 60  
QY 61 VLENLLVLAITSHMRSRRMYVCLVNTTSDLTGAYLANVLSGARFRLAPAOQWL 120  
DB 61 VLENLLVLAITSHMRSRRMYVCLVNTTSDLTGAYLANVLSGARFRLAPAOQWL 120  
QY 121 REGLEFALAASTFSLFTAGGERFATWVPVASEGATKTSRVYIGICWLLAALGMLP 180  
DB 121 REGLEFALAASTFSLFTAGGERFATWVPVASEGATKTSRVYIGICWLLAALGMLP 180  
QY 181 LIGWNCICAFDRCSLLPIYSKRYILFCLVIFAGVLAITMGLYGAIFRLVQASQKAPRP 240  
DB 181 LIGWNCICAFDRCSLLPIYSKRYILFCLVIFAGVLAITMGLYGAIFRLVQASQKAPRP 240  
QY 241 AARRKARRLKTYMLLAFIVCWGPFLEGLLADVFGSNLMAOEYLRGMIMLALAVLNS 300  
DB 241 AARRKARRLKTYMLLAFIVCWGPFLEGLLADVFGSNLMAOEYLRGMIMLALAVLNS 300  
QY 301 AVNPITTSFRSREVCRAVLSFLCCGCLRLGMRGDCCLARAVEAHSASTTDSLRPRDS 360  
DB 301 AVNPITTSFRSREVCRAVLSFLCCGCLRLGMRGDCCLARAVEAHSASTTDSLRPRDS 360  
QY 361 FRGSRSLSPMRREPLSSISVSRSI 384  
DB 361 FRGSRSLSPMRREPLSSISVSRSI 384

RESULT 4  
AAM48981 ID AAM48981 standard; Protein; 384 AA.  
AC AAM48981;  
XX  
DT 26-APR-2002 (first entry)  
XX Human endothelial differentiation, G-protein coupled receptor 6.  
XX  
KW Human: endothelial differentiation, G-protein coupled receptor 6;  
KW EDG6; haplotype; cancer; angiogenesis; inflammation; chromosome 19p13.3;  
KM cytostatic; antiinflammatory; gene therapy; SNP;  
KM single nucleotide polymorphism.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 155 /note= "optionally Arg"  
FT Misc-difference 164 /note= "optionally Ser"  
FT

FT Misc-difference 198 /note= "optionally Ser"  
FT Misc-difference 243 /note= "optionally Cys"  
FT Misc-difference 365 /note= "optionally Leu"  
FT Misc-difference 381 /note= "optionally Met"  
XX  
XX W0200206446-A2.  
XX 24-JAN-2002.  
XX  
XX 17-JUL-2001; 2001WO-US22523.  
XX  
XX 17-JUL-2000; 2000US-218727P.  
XX  
XX (GENA-) GENNISSANCE PHARM INC.  
XX  
XX Kilem SE, Koshy B;  
XX  
XX WPI; 2002-171804/22.  
XX N-PSDB; ABL45808, ABL45809.  
XX  
XX New genetic variants of endothelial differentiation, G-protein coupled  
XX receptor-6 gene for studying expression, function of the gene and  
XX expressing EDG6 protein for use in screening drugs to treat cancer,  
XX inflammation -  
XX  
XX Claim 28; Fig 3; 11pp; English.

CC The present invention provides the gene, protein and cDNA sequences of  
CC the human endothelial differentiation, G-protein coupled receptor 6  
CC (EDG6). Also identified are single nucleotide polymorphisms (SNPs) found  
CC within the sequences. The sequences can be used in the identification of  
CC the haplotype of an individual, and in the treatment of cancer,  
CC angiogenesis and inflammation. The present sequence is the EDG6 protein,  
CC the gene for which is found on chromosome 19p13.3.

CC  
SQ Sequence 384 AA:

Query Match 100.0%; Score 1966; DB 23; Length 384;  
Best Local Similarity 100.0%; Pred. No. 2.4e-206;  
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNATGTPVAPESCOQLAAGHSRLIVLHYNHSGRLAGRGPEDEGLALRGLSYAASCLV 60  
DB 1 MNATGTPVAPESCOQLAAGHSRLIVLHYNHSGRLAGRGPEDEGLALRGLSYAASCLV 60  
QY 61 VLENLLVLAITSHMRSRRMYVCLVNTTSDLTGAYLANVLSGARFRLAPAOQWL 120  
DB 61 VLENLLVLAITSHMRSRRMYVCLVNTTSDLTGAYLANVLSGARFRLAPAOQWL 120  
QY 121 REGLEFALAASTFSLFTAGGERFATWVPVASEGATKTSRVYIGICWLLAALGMLP 180  
DB 121 REGLEFALAASTFSLFTAGGERFATWVPVASEGATKTSRVYIGICWLLAALGMLP 180  
QY 181 LIGWNCICAFDRCSLLPIYSKRYILFCLVIFAGVLAITMGLYGAIFRLVQASQKAPRP 240  
DB 181 LIGWNCICAFDRCSLLPIYSKRYILFCLVIFAGVLAITMGLYGAIFRLVQASQKAPRP 240  
QY 241 AARRKARRLKTYMLLAFIVCWGPFLEGLLADVFGSNLMAOEYLRGMIMLALAVLNS 300  
DB 241 AARRKARRLKTYMLLAFIVCWGPFLEGLLADVFGSNLMAOEYLRGMIMLALAVLNS 300  
QY 301 AVNPITTSFRSREVCRAVLSFLCCGCLRLGMRGDCCLARAVEAHSASTTDSLRPRDS 360  
DB 301 AVNPITTSFRSREVCRAVLSFLCCGCLRLGMRGDCCLARAVEAHSASTTDSLRPRDS 360  
QY 361 FRGSRSLSPMRREPLSSISVSRSI 384  
DB 361 FRGSRSLSPMRREPLSSISVSRSI 384



RESULT 5  
 ID AAY06412  
 XX AAY06412 standard; Protein: 384 AA.  
 AC AAY06412;  
 DT 20-SEP-1999 (first entry)  
 XX  
 DE Human EDG-7 receptor homologue.  
 XX  
 EDG-7; human; HEDG-7; G protein coupled receptor; asthma;  
 adult respiratory distress syndrome; rheumatoid arthritis;  
 cardiac ischaemia; acute pancreatitis; septic shock; psoriasis;  
 acute cyclosporin nephrotoxicity; diabetic glomerulopathy;  
 lung damage; diagnosis; therapy; drug screening;  
 antiinflammatory.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Location/Qualifiers  
 FT 50..70  
 FT /note= "transmembrane domain I"  
 FT 71..81  
 FT /note= "intracellular domain IL-I"  
 FT 82..105  
 FT /note= "transmembrane domain II"  
 FT 106..124  
 FT /note= "extracellular domain EL-I"  
 FT 125..143  
 FT /note= "transmembrane domain IIT"  
 FT 144..163  
 FT /note= "intracellular domain IL-IIT"  
 FT 164..182  
 FT /note= "transmembrane domain IV"  
 FT 183..199  
 FT /note= "extracellular domain EL-IIT"  
 FT 200..224  
 FT /note= "transmembrane domain V"  
 FT 225..250  
 FT /note= "intracellular domain IL-IIT"  
 FT 251..272  
 FT /note= "transmembrane domain VI"  
 FT 273..285  
 FT /note= "extracellular domain EL-IIT"  
 FT 286..304  
 FT /note= "transmembrane domain VII"  
 FT 2  
 FT /note= "N-glycosylated"  
 FT 30  
 FT /note= "N-glycosylated"  
 FT 77  
 FT /note= "O-phosphorylated"  
 FT 159  
 FT /note= "O-phosphorylated"  
 FT 308  
 FT /note= "O-phosphorylated"  
 FT 360  
 FT /note= "O-phosphorylated"  
 FT 380  
 FT /note= "O-phosphorylated"  
 FT  
 XX MO935106-AZ.  
 XX  
 PD 15-JUL-1999.  
 XX  
 PF 30-DEC-1998; 98WC-CA01196.  
 XX  
 PR 30-DEC-1997; 97US-0070184.  
 XX  
 PA (ALIX ) ALLELIX BIOPHARMACEUTICALS INC.  
 XX  
 PI Gupta AK, Munroe DG, Zastawny RL;

XX  
 DR WPI; 1999-419322/35.  
 DR N-PSDB; AAX59367.  
 XX  
 PT A nucleic acid sequence encoding human EDG-7 receptor, useful for  
 PT treating disorders associated with aberrant expression  
 XX  
 PS Claim 12; Fig 2B; 72pp; English.  
 XX  
 CC The present sequence represents a novel human homologue of the  
 CC EDG-7 receptor, i.e. HEDG-7, a 7-transmembrane G protein coupled  
 CC receptor. The sequence is predicted from the cDNA insert of a  
 CC Jurkat T-cell clone (see AAX59367) and differs at only 2 amino acid  
 CC residues (140 and 178) from the sequence (see AAY06411) predicted  
 CC from BAC and PMC clones. Also claimed are an expression vector  
 CC comprising a HEDG-7 nucleotide sequence, a host cell, and methods  
 CC of using HEDG-7 receptor to identify potential HEDG-7 ligands and  
 CC antagonists. Ligands, agonists and antagonists of HEDG-7 may  
 CC be useful for detecting inflammation or disease associated with  
 CC abnormal levels of HEDG-7 expression. Detection of aberrant  
 CC expression of HEDG-7 can accelerate diagnosis and proper treatment  
 CC of abnormal conditions, e.g. adult respiratory distress, asthma,  
 CC rheumatoid arthritis, cardiac ischaemia, acute pancreatitis, septic  
 CC shock, psoriasis, acute cyclosporin nephrotoxicity and early  
 CC diabetic glomerulopathy, as well as lung damage following exposure  
 CC to cigarette smoke, asbestos or silica.  
 CC  
 XX  
 SQ Sequence 384 AA;  
 XX  
 Query Match 99.5%; Score 1956; DB 20; Length 384;  
 Best Local Similarity 99.5%; Pred. No. 3e-205;  
 Matches 382; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MNATGTPVAPESCCQQLAAGHSRLIVLHYNHSGRLAGRGEDGCGALNGLSYAASCLV 60  
 DB 1 MNATGTPVAPESCCQQLAAGHSRLIVLHYNHSGRLAGRGEDGCGALNGLSYAASCLV 60  
 QY 61 VLENILVLAATTSNMRSRWYVCLVNTTSDLTGAAVLANVLLSGARFRLAPAQWL 120  
 DB 61 VLENILVLAATTSNMRSRWYVCLVNTTSDLTGAAVLANVLLSGARFRLAPAQWL 120  
 QY 121 REGILFTALASTPSILFTAGREFATWVPVASEGATKTSRYVYIGLCWLLAALLGMLP 180  
 DB 121 REGILFTALASTPSILFTVGERPATWVPVASEGATKTSRYVYIFIDCWLLAALLGMLP 180  
 QY 181 LLGNWCLCAFDRCSLLPLYSKRYILFCLVTFAGVLAITMGLYGAIFRLVQASQKAPRP 240  
 DB 181 LLGNWCLCAFDRCSLLPLYSKRYILFCLVTFAGVLAITMGLYGAIFRLVQASQKAPRP 240  
 QY 241 AARRKARRLKTYIMTLAFLVCMGPLEGILLADYFGSNLMAQRYLGMOMIILAVLNS 300  
 DB 241 AARRKARRLKTYIMTLAFLVCMGPLEGILLADYFGSNLMAQRYLGMOMIILAVLNS 300  
 QY 301 AVNPITVSPRSREVCRAVLSFLCCGCLRLGMRGPDCLARAVERHSGASTTDSILRPDS 360  
 DB 301 AVNPITVSPRSREVCRAVLSFLCCGCLRLGMRGPDCLARAVERHSGASTTDSILRPDS 360  
 QY 361 FRGSRSLFRMRPEPLSSISVSRSI 384  
 DB 361 FRGSRSLFRMRPEPLSSISVSRSI 384  
 XX  
 RESULT 6  
 AAW90863  
 ID AAW90863 standard; Protein: 386 AA.  
 XX  
 AC AAW90863;  
 XX  
 DT 29-JUN-2000 (first entry)  
 XX  
 DE Murine edg6 protein.  
 XX  
 KW edg6; murine; G-coupled receptor; endothelial differentiation gene;

KM antiinflammatory; immunomodulatory; antimicrobial; antiallergic;  
 KM cytostatic; gene therapy; inflammation; autoimmune disease; allergy;  
 KM tumor; leukemia; lymphoma.  
 OS Mus sp.  
 PN DE19846979-A1.  
 XX 23-MAR-2000.  
 PD 13-OCT-1998; 98DE-1046979.  
 XX 11-SEP-1998; 98DE-1043240.  
 PR (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.  
 XX Graeler M, Bernhardt G, Lipp M;  
 PI MPI: 2000-258069/23.  
 DR N-PSDB; AAX82790.  
 XX  
 PT New human and murine G-coupled receptor EDG (endothelial  
 PT differentiation gene) 6, useful for modulating inflammatory and immune  
 PT reactions and for treatment of allergy or tumors -  
 PS Claim 2; Page 6; 12pp; German.  
 XX  
 CC This invention describes novel human and murine G-coupled receptors EDG  
 CC (endothelial differentiation gene) 6 (I and II). The products of the  
 CC invention have antiinflammatory, immunomodulatory, antimicrobial,  
 CC antiallergic and cytostatic activity. (I) and (II) are involved in signal  
 CC transduction. (I), (II) and their fragments, variants and mutants or  
 CC binding partners, are used therapeutically to modulate the function of  
 CC blood and body cells, particularly for inhibition of acute and chronic  
 CC inflammation and to raise specific antibodies against them. They are used  
 CC as a source of diagnostic oligonucleotides and for gene therapy.  
 CC Antibodies against (I) and (II) are useful for diagnosis and optionally  
 CC when coupled to therapeutic agents, toxins or other antibodies, to  
 CC modulate immune and inflammatory responses for example immunological  
 CC defects such as inflammation, infection, autoimmune diseases, allergy,  
 CC tumors, leukemia and lymphoma. This sequence represents the murine EDG6  
 CC protein described in the method of the invention.  
 XX  
 SQ Sequence 386 AA;  
 Query Match 82.6%; Score 1623; DB 21; Length 386;  
 Best Local Similarity 82.8%; Pred. No. 8.6e-169;  
 Matches 314; Conservative 28; Mismatches 35; Indels 2; Gaps 2;  
 QY 6 TPVAPESCQQLAAGSHSLVLYHNHSGRLAGRGPEP-GGLGALRGLSVAASCLVVEN 64  
 DB 8 TLVTPESCHRLAASGHSLLVLYHNHSGRLASRGSEDEGGLGMLRGPVAACTLVVEN 67  
 QY 65 LVLVAITSHHRSRRWYVYCLVNTTSLDLTGAAYLANVLSGARTFLAPQWFLREG 124  
 DB 68 AMVLAIAIYHRSRRWYVYCLVNTTSLDLTGLAYVAVNLISGRTTQLSVHMFLEGL 127  
 QY 125 LFTALASTFSLFTAGGERFATWVRVAESGATTSRYGFTIGLCWLLAALGLMPLIGW 184  
 DB 128 LFMALASTFSLFTAGGERFATWVR-VAESGATTSRYGFTIGLCWLLAALGLMPLIGW 186  
 QY 185 NCLTAFORCSSLLPLYSKRYITLFLVYFAGVLAITMGLYGAIFRVLVQASQKAPRPAR 244  
 DB 187 NCVCAPFPCSSLLPLYSKRYITLFLVYFAGVLAITMGLYGAIFRVLVQASQKAPRPAR 246  
 QY 245 KARRLTVMITLAFVLCVCGPFLGILLADVFSGNLMAOYELGMDITLALVANSVNP 304  
 DB 247 KSRRLMTVMITLAFVLCVCGPFLGILLADVFSGNLMAOYELGMDITLALVANSVNP 306  
 QY 305 IYFSRREVCRAVLSFLCGCCLNLRGPGDCLARAVEAHSGASTTDSLSLRPDSFRGS 364  
 DB 307 LIYFSRREVCRAVLSFLCGCCLNLRGPGDCLARAVEAHSGASTTDSLSLRPDSFRGS 366

OY 365 RSLSRMRPLSSISSVRS 383  
 DB 367 RSLSRMRPLSSISSVRS 385

RESULT 7  
 AAW80955  
 ID AAW80955 standard; Protein: 509 AA.  
 AC AAW80955;  
 XX 16-FEB-1999 (first entry)  
 DE Amino acid sequence of an edg-1 receptor.  
 XX  
 DE NEDG; edg-1-like receptor; stimulation; cell proliferation; cancer;  
 KM genetic defect; sickle cell anaemia; agonist; antagonist; antibody;  
 KW probe.  
 XX Homo sapiens.  
 OS

Key location/Qualifiers  
 FT Misc-difference 20 /note= "encoded by NGG"  
 FT Misc-difference 26 /note= "encoded by GNT"  
 XX  
 XX WC9848016-A1.  
 XX 29-OCT-1998.  
 XX 17-APR-1998; 98WO-US07830.  
 XX 24-APR-1997; 97US-0845566.  
 XX (INCY-) INCYTE PHARM INC.  
 XX Au-Young J, Guegler KJ;  
 DR MPI: 1998-609989/51.  
 DR N-PSDB; AAW68602.  
 XX

PT New Edg-1 like receptor, NEDG - useful, e.g. to stimulate cell  
 PT proliferation and screen for antagonists to treat disorders of  
 PT excessive cell proliferation such as cancers  
 PS Claim 1; Fig 1A-1E; p 63pp; English.  
 XX

This is the amino acid sequence of an edg-1-like receptor (NEDG), used  
 CC in the method of the invention. NEDG can be used to stimulate cell  
 CC proliferation, especially by expressing encoding polynucleotides,  
 CC either in vivo (e.g. to promote cell regeneration/differentiation) or  
 CC in vitro (e.g. to stimulate cell proliferation for transplantation).  
 CC Such administration may be useful therapeutically, e.g. to stimulate  
 CC proliferation of cells selected for their ability to fight an  
 CC infection or a cancer, or correct a genetic defect, e.g. sickle cell  
 CC anaemia. It can be used to screen for agonists (useful to  
 CC treat/prevent disorders of abnormal cell growth and differentiation  
 CC as above) and antagonists, and to generate antibodies. The antagonists  
 CC can be combined with a suitable carrier in pharmaceutical compositions,  
 CC useful to treat/prevent disorders of excessive cell proliferation,  
 CC especially cancers. The NEDG-specific antibodies may be used directly  
 CC as antagonists, or indirectly as a targeting or delivery mechanism  
 CC to bring pharmaceutical agents to NEDG-expressing cells. They are  
 CC also useful to diagnose conditions or diseases characterised by NEDG  
 CC expression, and to monitor therapeutic interventions. They are useful  
 CC to produce antisense sequences for therapeutic administration to  
 CC modulate/prevent NEDG expression. They may also be used to produce  
 CC probes useful to detect related sequences, or for gene mapping.

SQ Sequence 509 AA;

Query Match 73.1%; Score 1437; DB 19; Length 509;

Best Local Similarity 76.3%, Pred. No. 2.9e-148;  
Matches 303; Conservative 8; Mismatches 42; Indels 44; Gaps 6;

```

OY 1 MNAGTTPVAPESCOQLAAGHSRLIVLHYNSGRLAGRGSPEDGGLRLGLSVASCLV 60
DB 1 MNAGTTPVAPESCOQLAAGHSRLIXLHYNSGRLAGRGSPEDGGLRLGLSVASCLV 60
OY 61 VLENILVLAATSHMRSRMWYVCLVNTITSDLLTGAAVLAVLLSGARTFLAPAOFWL 120
DB 61 VLENILVLAATSHMRSRMWYVCLVNTITSDLLTGAAVLAVLLSGARTFLAPAOFWL 120
OY 121 REGILFTALAASTFSLFTAGRPFATWVRPVAESGATKTSRYVGFGLCWLIALGLMLP 180
DB 121 REGILFTALAASTFSLFTAGRPFATWVRPVAESGATKTSRYVGFGLCWLIALGLMLP 180
OY 181 ILGNMCCARDRCSSILPLXSKRYTLPCLYTFAGVLAATMGXYGAFRLVQASQKAPRP 240
DB 181 ILGNMCCARDRCSSILPLXSKRYTLPCLYTFAGVLAATMGXYGAFRLVQASQKAPRP 240
OY 241 AARRKARRLKTVMILLLAFVYCGPFLFGLLADVRGSSNMAQ----- 283
DB 241 AARRKARRLKTVMILLLAFVYCGPFLFGLLADVRGSSNMAQ----- 283
OY 284 -EYLGMDWILALAVLNSAVNPDIYSFRSREVCRAVLSFLCCGCLRLQMRGPDCCLARAV 342
DB 297 RPOGLGCPHHLPLPOOGGYOS-----RAOPSSAAAGVSGWAC-----EGPGDCLARAV 343
OY 343 EAHSGASTTDSLSLRPDSFRGSRSLSPFRMREPLSSIS 379
DB 344 EAHSGA-----SNHROLSEAKGOLS-----RLPLAQLS 371

```

## RESULT 8

AAM81577  
ID AAM81577 standard; Protein; 276 AA.

```

AC AAM81577;
DE 22-FEB-1999 (first entry)
XX EDG-1-like G-protein coupled receptor polypeptide.
XX EBV-induced G-protein coupled receptor; EB1-2; Epstein-Barr Virus; ulcer;
XX endothelium-differentiation gene; EDG-1-like G-protein coupled receptor;
XX recombinant; agonist; asthma; Parkinson's disease; heart failure; asthma;
XX hypertension; urinary retention; osteoporosis; allergy; psychosis; depression;
XX angina pectoris; myocardial infarction; migraine; stroke; cancer;
XX migraine; vomiting; stroke; eating disorder; migraine headache; cancer;
XX prostatic hypertrophy; detection; drug screening.
XX Homo sapiens.
OS
XX
FH Key
FH Peptide 1..16 Location/Qualifiers
FT /note= "signal peptide"
FT Protein 17..276
FT /note= "mature protein"
FT Misc-difference 207 /note= "encoded by GAC"
FT Misc-difference 218 /note= "encoded by AAG"
FT Misc-difference 220 /note= "encoded by TAC"

```

```

XX MO9850549-A2.
XX 12-NOV-1998.
XX 07-MAY-1998; 98WO-0509048.
XX 07-MAY-1997; 97US-0852824.
XX (HUMA-) HUMAN GENOME SCI INC.

```

```

XX LI Y, Ruben SM;
XX WPI: 1999-034722/03.
XX N-PSDB: AAV69761.
XX New isolated human G-protein coupled receptors - used to develop
XX products for treating e.g. asthma, Parkinson's disease, heart
XX failure, osteoporosis, hypertension, psychoses, eating disorders or
XX cancers
XX Claim 1; Fig 3A-B; 65bp; English.
XX
XX This represents a EDG-1-like G-protein coupled receptor polypeptide.
XX The encoding DNA is deposited under the accession number ATCC No:
XX 209004. The invention provides two human G-protein coupled receptor
XX polypeptides. The polypeptides are human Epstein-Barr Virus (EBV)-induced
XX G-protein coupled receptor, designated EB1-2 polypeptide and a human
XX endothelium-differentiation gene (EDG) like G-protein coupled receptor,
XX designated EDG-1-like G-protein coupled receptor. Vectors comprising the
XX EB1-2 and EDG-1-like G-protein coupled encoding DNA can be used to transform
XX host cells for the recombinant production of the proteins. Agonists for
XX G-protein coupled receptors can be used for the treatment of asthma,
XX Parkinson's disease, acute heart failure, hypertension, urinary retention
XX and osteoporosis. Antagonists can be used for the treatment of
XX hypertension, angina pectoris, myocardial infarction, ulcers, asthma,
XX allergies, psychoses, depression, migraine, vomiting, stroke, eating
XX disorders, migraine headaches, cancer and benign prostatic hypertrophy.
XX The products can also be used for detection, diagnosis and drug
XX screening.
XX
XX Sequence 276 AA:

```

Query Match 65.2%, Score 1281; DB 20; Length 276;  
Best Local Similarity 97.3%, Pred. No. 1.5e-131;  
Matches 252; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 MNAGTTPVAPESCOQLAAGHSRLIVLHYNSGRLAGRGSPEDGGLRLGLSVASCLV 60
DB 17 MNAGTTPVAPESCOQLAAGHSRLIVLHYNSGRLAGRGSPEDGGLRLGLSVASCLV 76
OY 61 VLENILVLAATSHMRSRMWYVCLVNTITSDLLTGAAVLAVLLSGARTFLAPAOFWL 120
DB 77 VLENILVLAATSHMRSRMWYVCLVNTITSDLLTGAAVLAVLLSGARTFLAPAOFWL 136
OY 121 REGILFTALAASTFSLFTAGRPFATWVRPVAESGATKTSRYVGFGLCWLIALGLMLP 180
DB 137 REGILFTALAASTFSLFTAGRPFATWVRPVAESGATKTSRYVGFGLCWLIALGLMLP 196
OY 181 ILGNMCCARDRCSSILPLXSKRYTLPCLYTFAGVLAATMGXYGAFRLVQASQKAPRP 240
DB 197 ILGNMCCARDRCSSILPLXSKRYTLPCLYTFAGVLAATMGXYGAFRLVQASQKAPRP 256
OY 241 AARRKARRLKTVMILLLAFVYCGPFLFGLLADVRGSSNMAQ 259
DB 257 AARRKARRLKTVMILLLAFVYCGPFLFGLLADVRGSSNMAQ 275

```

## RESULT 9

ABB04481  
ID ABB04481 standard; Protein; 348 AA.

```

AC ABB04481;
DE 15-MAR-2002 (first entry)
XX Human Lyso-1 receptor.
XX Human, Lyso-1 receptor; G-protein coupled receptor; chromosome 11p;
XX EDG6 receptor; Gessler-Wilm's tumour.
XX Homo sapiens.
XX

```



```
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KM tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KM diagnostics; forensic test; gene mapping; genetic disorder;
XX biodiversity; gene therapy; nutrition.
OS Homo sapiens.
XX
XX WO200154477-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US02687.
XX
XX 25-JAN-2000; 2000US-0491404.
XX 17-JUL-2000; 2000US-0617746.
XX 03-AUG-2000; 2000US-0631451.
XX 15-SEP-2000; 2000US-0663870.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX
XX WPI: 2001-476164/51.
XX N-PSDB: AAH98657.
XX
XX Isolated polypeptide for treatment of diseases, diagnostics, raising
XX antibodies and research use -
XX
XX Claim 20; Page 1051; 1275pp; English.
XX
XX The present invention provides the protein and coding sequences of novel
XX proteins from a variety of organisms, including human, dog, cat, horse,
XX cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
XX urchin and tomato. These were derived from expressed sequence tags (ESTs)
XX from the organism of interest. They can be used in diagnostics,
XX forensics, gene mapping, identification of mutations, to assess
XX biodiversity and for nutritional purposes. The present sequence is a
XX protein of the invention.
XX
XX Sequence 130 AA:
XX
XX Query Match 34.1%; Score 671; DB 22; Length 130;
XX Best Local Similarity 100.0%; Pred. No. 3.2e-65;
XX Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 255 MILAFLVCMGPIFGILLADVFGSNLMAOEYLKGMIMILAVLNSAVNPITISFRRSREV 314
XX |||||
XX 1 MILAFLVCMGPIFGILLADVFGSNLMAOEYLKGMIMILAVLNSAVNPITISFRRSREV 60
XX
XX 315 CRAVLFLCCGCLRLGMRGPGDCLARAVEAHSASTTDSLSLRPDSFRSRLSFFMRKP 374
XX |||||
XX 61 CRAVLFLCCGCLRLGMRGPGDCLARAVEAHSASTTDSLSLRPDSFRSRLSFFMRKP 120
XX
XX 375 LSSISVSRSI 384
XX |||||
XX 121 LSSISVSRSI 130
XX
XX
XX RESULT 12
XX AAM01664
XX ID AAM01664 standard; Protein: 383 AA.
XX
XX AAM01664;
XX
XX 01-APR-1997 (first entry)
XX
XX p(rat-edg), G-protein coupled receptor.
XX
XX p(H218); G-protein coupled receptor; cell differentiation; proliferation;
XX proline directed kinase; cell division; growth factor response; rat-edg;
XX therapy; diagnosis.
```

```
XX OS Rattus rattus.
XX
XX US5585476-A.
XX
XX 17-DEC-1996.
XX
XX 15-FEB-1994; 94US-0196989.
XX
XX 15-FEB-1994; 94US-0196989.
XX
XX (MACL/) MACLENNAN A J.
XX
XX Macleennan AJ;
XX
XX WPI: 1997-051235/05.
XX N-PSDB: AAT58506.
XX
XX DNA encoding rat protein p(H218) - associated with cell
XX proliferation and/or differentiation
XX
XX Example 9; Column 17-20; 33pp; English.
XX
XX This sequence is a novel rat protein p(rat-edg), a member of the
XX G-protein coupled receptor (GPR) superfamily of proteins. The
XX amino acid similarity between p(H218) (AAM01663) and p(rat-edg) suggests
XX that they may be activated by the same endogenous ligand(s). The
XX expression pattern of mRNA transcripts of both genes in cell lines,
XX various rat tissues and developing rat brain suggests that they both
XX play a role in cell proliferation and/or differentiation.
XX
XX Sequence 383 AA:
XX
XX Query Match 33.9%; Score 666; DB 18; Length 383;
XX Best Local Similarity 40.0%; Pred. No. 4.9e-64;
XX Matches 150; Conservative 69; Mismatches 126; Indels 30; Gaps 8;
XX
XX 1 MNATGTPAPAPSCCOOLAAGHSRLVLVHNHSGRLAGRGPGEDGALRGLSVNASCIV 60
XX |||||
XX 2 VSTSIPIYVKALRQGVSDYGNVDIIVRYNTGKL-NIGVEKDHGIKLTSVFILICCLI 60
XX
XX 61 VLENLVLAAITSHMSRRWYVCIVNTTSLDLTGAAVLANVLGAKTFRPAPQWFL 120
XX |||||
XX 61 ILENFVLLTITWKTKFRRPMYFIGNLALSDLLAGAVATNLLLSGATYKLPAPQWFL 120
XX
XX 121 REGLLFTALASTPSLFTTAGEFRATWVRPAESGATTSVYFGICWMLLALGLMP 180
XX |||||
XX 121 REGSMFVALSASFSLAIAIERYITMLKMLHNG-SNHSRSPILLISACWYISLILGLP 179
XX
XX 181 ILGNMCLAFDRCSLLPLYSKRYILFCLVIFAGVATIMGLYGAIFRLVQASQ----- 235
XX |||||
XX 180 IMGWNCISSLSGCSVLPDLKHKHYILFCTYVFTLLLSIVLYCRISVLTFRSRLTFR 239
XX
XX 236 ---KAPPARARRKRLIKYVLMILAFVLCWGPFLCLLADYFGSNLMAOEYLKGMIM 291
XX |||||
XX 240 KNISKASR--SSSEKSLALIKTVIIVLSVFIACWMPFLTLILDV-GCKAKTCDILYAEY 296
XX
XX 292 ILALAVNSAVNPITISFRRSREVCRVAVLSPL-CCGCLRLGMRG-----PGDCLARAVE 343
XX |||||
XX 297 FLVLAVLNSGTPITITLTNKMERRAFIRIISCCCPNGDSAGKAFKRIITGMEFSR--- 353
XX
XX 344 AHSASTTDSLSLRP 358
XX |||||
XX 354 -----SKSDNSHPQ 363
XX
XX
XX RESULT 13
XX AAW87791
XX ID AAW87791 standard; Protein: 383 AA.
XX
XX AAW87791;
XX
XX 12-MAR-1999 (first entry)
```

XX Rat-edg, G-protein coupled receptor superfamily member.  
DE  
XX H218: G-protein coupled receptor superfamily; cell proliferation;  
KW cell differentiation; cancer; rat-edg; diagnosis; therapy.  
XX  
XX Rattus sp.  
OS  
XX US5856443-A.  
PN  
XX 05-JAN-1999.  
PD  
XX 06-DEC-1996; 96US-0760936.  
PF  
XX 15-FEB-1994; 94US-0196989.  
PR 06-DEC-1996; 96US-0760936.  
XX  
PA (MACL/) MACLENNAN A J.  
PI  
XX MacLennan AJ;  
PI  
XX WPI: 1999-105192/09.  
DR N-PSDB; AAV84036.  
DR  
XX  
XX New isolated polynucleotide encoding a G-protein coupled receptor -  
PT used to develop products for use in diagnosis and therapy of  
PT conditions involving cell proliferation and differentiation  
XX  
XX Disclosure; Columns 23-26; 33pp; English.  
PS  
XX  
XX The present sequence represents a protein designated rat-edg. The rat-edg  
CC polypeptide is a member of the G-protein coupled receptor superfamily.  
CC The specification also describes a protein designated H218, which  
CC is also a member of the G-protein coupled receptor superfamily. H218  
CC is involved in cell proliferation and differentiation, and in disease  
CC states such as cancer. The polynucleotides and polypeptides can be  
CC used to develop products for diagnosis and therapy.  
XX  
XX Sequence 383 AA:  
SQ

Query Match 33.9%; Score 666; DB 20; Length 383;  
Best Local Similarity 40.0%; Pred. No. 4.9e-64;  
Matches 150; Conservative 69; Mismatches 126; Indels 30; Gaps 8;

QY 1 MNATGTPVAPESCQQLAAGHSRLIYLHNHSGRLAGRGPEDEGIGALRGLSVASCLV 60  
DB 2 VESTSIPVVKALRSQVSDYGNVDIIVRHNYTKL-NIGVEKDHGIKLSVFIILCCL 60  
QY 61 VLENLLVLAITSHMRSRWVYCLVNTLSDLTGAAYLANVLLSGARTFRLAPQWFL 120  
DB 61 ILENIFVLLTIWTKRHFHPRMYFFIGNLALSDLAGVAYTANLLSGATYKLPAPQWFL 120  
QY 121 REGILFTALASTFSLFTAGREFATMVRPVAESGATKTSRYVGTIGLCWLLAALLGMLP 180  
DB 121 RGSMPVALSASFSLAIAIRYITMLMKKLHNG-SNSRSFLLISACWVISLILGGLP 179  
QY 181 LIGWNCLEAFDRCSLLPLYSKRYILFCLVIFAGVLATIMGLYGAIFRLVOASQ----- 235  
DB 180 IMGWNCISLSSCTVLPYHKHYILFCTVETLLSLIVILKYCYSLVTRSRRLFR 239  
QY 236 -----KAPRAARRRARLLKTYMLLLAFYCWGFLFELLADYFGSLMAOEYLRGMDW 291  
DB 240 KNISKASR--SSEKSLALKTVIIVLSYFIACWAPLFTLLLDV-GCKAKTCDILYKAEY 296  
QY 292 ILALVLSAVNPITISFRSREVCRAVLSPL-CCGCLRLGMRG-----PGDCLARAVE 343  
DB 297 FLVLAVLSGNPIYITITLTKEMKRAPIRITISCCCKPWSGDSAGFKRPITIGMEFSR--- 353  
QY 344 AHGSASTTDSILRPR 358  
DB 354 -----SKSDNSHPQ 363

RESULT 14  
AAB28521  
ID AAB28521 standard; Protein; 382 AA.  
XX  
XX AAB28521;  
AC  
XX 06-FEB-2001 (first entry)  
DT  
XX  
XX Mouse EDG1 polypeptide.  
DE  
XX  
XX Mouse; EDG1; antimicrobial; cytostatic; analgesics; antidiabetic;  
KW antiparkinsonian; cardiant; osteopathic; anorectic; antiasmatic;  
KW antiemetic; tranquiliser; antiallergic; neuroleptic; antiulcer;  
KW antidepressant; infection; cancer; diabetes; eating disorder;  
KW asthma; Parkinson's disease; heart failure; hypotension; hypertension;  
KW osteoporosis; angina pectoris; stroke; ulcer; allergy;  
KW benign prostatic hypertrophy; migraine; vomiting; neurological disorder.  
XX  
XX Mus musculus.  
OS  
XX  
XX WO200059529-A1.  
PN  
XX 12-OCT-2000.  
PD  
XX  
XX 05-APR-2000; 2000MO-US09075.  
PF  
XX 05-APR-1999; 99US-0127696.  
PR 04-APR-2000; 2000US-0127696.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
PA  
PI  
XX Tsui P;  
PI  
XX WPI: 2000-664960/64.  
DR N-PSDB; AAC63385.  
DR  
XX  
XX Murine EDG1 polypeptide useful for treating infections and diseases  
PT such as cancer, diabetes, Parkinson's disease, allergies, and  
PT neurological disorders such as anxiety schizophrenia, depression -  
XX  
XX Claim 1; Page 31-32; 34pp; English.  
XX  
XX The present sequence is mouse EDG1. Agonists and antagonists of  
CC mouse EDG1 are useful for treating patients with enhanced  
CC or inhibited activity or expression of EDG1. The EDG1 polypeptide is  
CC useful for treating abnormal conditions such as infections caused by  
CC bacteria, fungi etc., pain, cancer, diabetes, obesity, anorexia, bulimia,  
CC asthma, Parkinson's disease, acute heart failure, hypotension,  
CC hypertension, urinary retention, osteoporosis, angina pectoris,  
CC myocardial infarction, stroke, ulcers, allergies, benign prostatic  
CC and neurological disorders including anxiety, schizophrenia, manic  
CC depression, depression, delirium, dementia and severe mental retardation  
CC and dyskinesias such as Huntington's disease or Gilles de la Tourette's  
CC syndrome.  
XX  
XX Sequence 382 AA:  
SQ

Query Match 33.2%; Score 653; DB 21; Length 382;  
Best Local Similarity 39.5%; Pred. No. 1.3e-62;  
Matches 148; Conservative 67; Mismatches 130; Indels 30; Gaps 8;

QY 1 MNATGTPVAPESCQQLAAGHSRLIYLHNHSGRLAGRGPEDEGIGALRGLSVASCLV 60  
DB 1 MVESTSIPVVKALRSVSDYGNVDIIVRHNYTKL-NIGAEKDHGIKLSVFIILCCLPI 59  
QY 61 VLENLLVLAITSHMRSRWVYCLVNTLSDLTGAAYLANVLLSGARTFRLAPQWFL 120  
DB 60 ILENIFVLLTIWTKRHFHPRMYFFIGNLALSDLAGVAYTANLLSGATYKLPAPQWFL 119  
QY 121 REGILFTALASTFSLFTAGREFATMVRPVAESGATKTSRYVGTIGLCWLLAALLGMLP 180  
DB 120 RGSMPVALSASFSLAIAIRYITMLMKKLHNG-SNSRSFLLISACWVISLILGGLP 178

OY 181 LIGNNCLAFEDRCSSLLPLYSKRYILFCVIFAGVLATIMGLYGAIFRLVQASGQ----- 235  
DB 179 IMGNCCSSLSGCTVPLVHKHHTILFCTVIFLLLLSTIYLKCRISYLRTRSRLLTR 238  
OY 236 ----KAPPPARRKARLLTVMILIAFLVCGPLFGLLADYFGSNIMAOEYLGM 291  
DB 239 KNISKASR--SSEKSLALITVITIVLVSFTACWAPLFLILLDV--GCKAKTCIDILYKAY 295  
OY 292 ILALAVINSAVNPITYFSRSEVCRAVLSFL--CCGCLRLGMRG-----PGDCLARAWE 343  
DB 296 FLVLAIVNSGTNPITYITLTKEMKRAIRIVSCCKPCPDGSDAKKFKRPIIPGMEFSR--- 352  
OY 344 AHSASSTTDSLSLR 358  
DB 353 ----SKSDNSSHPQ 362

## RESULT 15

AA71305  
ID AA71305 standard; Protein; 500 AA.

XX  
AC AA71305;

DT 02-NOV-2000 (first entry)

DE Human orphan G protein-coupled receptor hCHN6.

KW Human; orphan G protein-coupled receptor; GPCR; hCHN6; drug screening;  
transmembrane receptor; expressed sequence tag; EST; signal cascade.

XX Homo sapiens.

OS  
PN WO200031258-A2.

PD 02-JUN-2000.

PF 13-OCT-1999; 99WO-US23687.

XX 20-NOV-1998; 98US-0109213.  
PR 16-FEB-1999; 99US-0120416.  
PR 26-FEB-1999; 99US-0121852.  
PR 12-MAR-1999; 99US-0123946.  
PR 12-MAR-1999; 99US-0123949.  
PR 28-MAY-1999; 99US-0136436.  
PR 28-MAY-1999; 99US-0136437.  
PR 28-MAY-1999; 99US-0136439.  
PR 28-MAY-1999; 99US-0136567.  
PR 28-MAY-1999; 99US-0137127.  
PR 28-MAY-1999; 99US-0137131.  
PR 29-JUN-1999; 99US-0141448.  
PR 29-SEP-1999; 99US-0156555.  
PR 29-SEP-1999; 99US-0156633.  
PR 29-SEP-1999; 99US-0156634.  
PR 29-SEP-1999; 99US-0156653.  
PR 01-OCT-1999; 99US-0157280.  
PR 01-OCT-1999; 99US-0157281.  
PR 01-OCT-1999; 99US-0157282.  
PR 01-OCT-1999; 99US-0157293.  
PR 01-OCT-1999; 99US-0157294.  
PR 12-OCT-1999; 99US-0416760.  
PR 12-OCT-1999; 99US-0417044.

XX (AREN-) ARENA PHARM INC.

XX Chen R, Dang HT, Liaw CW, Lin I;

XX WPI; 2000-400068/34.

XX N-PSDB; AAD01132.

XX Novel human orphan G protein-coupled receptors and the encoding cDNAs  
XX for use in the identification of G protein-coupled receptor agonists -

PS Claim 58; Page 79-81; 102pp; English.

XX The present amino acid sequence is the hCHN6, an endogenous human  
CC orphan G protein-coupled receptor (GPCR). The hCHN6 cDNA was identified  
CC using ESTs (expressed sequence tag) AA804531 and 2134670 as a probe.  
CC The orphan GPCR of the invention, like all GPCRs has seven transmembrane  
CC alpha helices with an extracellular N-terminus and an intracellular  
CC C-terminus. However, no endogenous ligands has yet been identified for  
CC the proteins of the invention. The orphan GPCRs may be used in the  
CC identification of their endogenous ligands, and to screen potential GPCR  
CC agonists and antagonists for use as pharmaceutical agents. The proteins  
CC may also be used in the study of GPCR-mediated signalling cascades, and  
CC to elucidate their precise role in normal and diseased human conditions.  
CC Nucleic acid encoding human orphan GPCRs may be used for tissue  
CC localisation expression analysis to provide information about their  
CC function in healthy and pathological states.

SO Sequence 500 AA:

Query Match 32.98; Score 647.5; DB 21; Length 500;

Best local similarity 40.8%; Pred. No. 7.5e-62;

Matches 154; Conservative 50; Mismatches 126; Indels 47; Gaps 7;

OY 22 SRLVLYHNHSGRLAGSGPEDEGLGRLRGSLVAASCIIVLENTLYLAITSHMRSRMV 81  
DB 114 SEVTLVHNHTGKLRGASTPGAGLRADAVCYLAVCFIVLEAVLVLGRRHPPHAPM 173  
OY 82 YYCLVNTITLSDLTGAAYLANVLISGARFFRLAPACMFUREGILFTALASTFSLFTAG 141  
DB 174 FLILGSLTSLDLAGAAYANILISGFLTKLSPALMFAREGVEFALVASTYLALIAL 233  
OY 142 ERFATMVR---PVAESGATKTSRVYIGLCMLLAALIGMLPLIGNNCLCAFDDCSSLL 197  
DB 234 ERSLTMARCGPAPVSSRG---RTLMAAAAGVSLILGLPLALGWNCLGRIDKCYTL 288  
OY 198 PLYSKRYILFCVIFAGVLATIMGLYGAIFRLVQASGQAR-RPA-----ARRAR- 247  
DB 289 PLVAKAVVLFVLAFAVGIILALCALYARIYCYRANARRLPARPAGTSTRARRKPRS 348  
OY 248 -RLKTYVIMILIAFLVCGPLFGLLADYFGSNIMAOEYLKGMILALAVINSAVNPIT 306  
DB 349 LALRLTSLVLAFAVACGPLFGLLIDVACPARTCPVLLDA-DPLGLAMANSLLNPIT 407  
OY 307 YSFRSEVRCRAVLSFLCCG-----CLRIGMRGPDCIAR 340  
DB 408 YTLNRPDLRHALLRLVCCGRHSCGRDPSGSQASAAEASGILRCLPPGLDGSPSGSER 467  
OY 341 AVEAHSASSTTDSLSLR 357  
DB 468 SSPQRDGLDTSSTGSP 484

Search completed: December 9, 2002, 12:04:10  
Job time : 40 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 12:03:07 : Search time 35 Seconds  
(Without alignments)  
2260.631 Million cell updates/sec

Title: US-09-786-926-1  
Perfect score: 1966  
Sequence: 1 MNATGVPAPESCQQLAAG.....RSLSPFMRPELSSISVRSI 384

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description          |
|------------|-------|-------------|--------|-----------|----------------------|
| 1          | 1966  | 100.0       | 384    | 4 095977  | 095977 homo sapien   |
| 2          | 1623  | 82.6        | 386    | 11 0920L1 | 0920L1 mus musculu   |
| 3          | 655.5 | 33.3        | 362    | 13 09DDK4 | 09ddk4 brachydanio   |
| 4          | 653   | 33.2        | 382    | 11 09DC35 | 09dc35 mus musculu   |
| 5          | 649   | 33.0        | 382    | 11 09R235 | 09r235 mus musculu   |
| 6          | 645.5 | 32.8        | 398    | 4 09H228  | 09h228 homo sapien   |
| 7          | 643.5 | 32.7        | 378    | 11 09P0U9 | 09p0u9 mus musculu   |
| 8          | 640   | 32.6        | 384    | 13 09PU08 | 09pu08 fugu rubrip   |
| 9          | 638   | 32.5        | 400    | 11 09UKM5 | 09ukm5 rattus norv   |
| 10         | 636   | 32.3        | 400    | 11 09IX56 | 09ix56 mus musculu   |
| 11         | 635.5 | 32.3        | 382    | 4 09NYN8  | 09ny8 homo sapien    |
| 12         | 632   | 32.1        | 387    | 11 099NM8 | 099nm8 mus musculu   |
| 13         | 628   | 31.9        | 400    | 11 09QY79 | 09qy79 rattus norv   |
| 14         | 618   | 31.4        | 326    | 11 099NR2 | 099nr2 dipodomys h   |
| 15         | 614   | 31.2        | 326    | 11 099NR4 | 099nr4 hystrix bra   |
| 16         | 613.5 | 31.2        | 326    | 11 099NR1 | 099nr1 heterocephala |

|    |       |      |     |           |                     |
|----|-------|------|-----|-----------|---------------------|
| 17 | 613   | 31.2 | 326 | 6 09BF63  | 09bf63 ochotona hy  |
| 18 | 612.5 | 31.2 | 326 | 11 099NR5 | 099nr5 mus musculu  |
| 19 | 612.5 | 31.2 | 326 | 11 099NR3 | 099nr3 erethizon d  |
| 20 | 611.5 | 31.1 | 326 | 6 09BF60  | 09bf60 lemur catia  |
| 21 | 610.5 | 31.1 | 326 | 11 099NR0 | 099nr0 cavia tachu  |
| 22 | 610   | 31.0 | 326 | 6 09BF73  | 09bf73 condylura c  |
| 23 | 609   | 31.0 | 326 | 6 09BF66  | 09bf66 elephanulu   |
| 24 | 609   | 31.0 | 326 | 11 099NR7 | 099nr7 agouti lacz  |
| 25 | 608   | 30.9 | 326 | 6 09BF70  | 09bf70 trichechus   |
| 26 | 608   | 30.9 | 326 | 6 09BF57  | 09bf57 hylobates c  |
| 27 | 608   | 30.9 | 326 | 6 09BF49  | 09bf49 tiragelaphus |
| 28 | 608   | 30.9 | 326 | 6 09BF48  | 09bf48 sus scrofa   |
| 29 | 608   | 30.9 | 326 | 6 09BF47  | 09bf47 okapia john  |
| 30 | 607.5 | 30.9 | 326 | 11 099NR9 | 099nr9 hydrochoeru  |
| 31 | 607   | 30.9 | 326 | 6 09BF58  | 09bf58 macaca mula  |
| 32 | 607   | 30.9 | 326 | 11 099NR8 | 099nr8 castor cana  |
| 33 | 606.5 | 30.8 | 326 | 6 09BF53  | 09bf53 rousettus l  |
| 34 | 606   | 30.8 | 325 | 11 099NR9 | 099nr9 tamias stri  |
| 35 | 606   | 30.8 | 326 | 6 09BF72  | 09bf72 sorex arane  |
| 36 | 606   | 30.8 | 326 | 6 09BF67  | 09bf67 macrosellid  |
| 37 | 605   | 30.8 | 326 | 6 09BF68  | 09bf68 loxodonta a  |
| 38 | 605   | 30.8 | 326 | 6 09BF46  | 09bf46 equus cabal  |
| 39 | 604   | 30.7 | 326 | 6 09BF75  | 09bf75 myrmecophag  |
| 40 | 604   | 30.7 | 326 | 6 09BF74  | 09bf74 erinaceus c  |
| 41 | 604   | 30.7 | 326 | 6 09BF69  | 09bf69 procavia ca  |
| 42 | 604   | 30.7 | 326 | 6 09BF52  | 09bf52 megaptera n  |
| 43 | 604   | 30.7 | 326 | 6 09BF45  | 09bf45 ceratotheri  |
| 44 | 602.5 | 30.6 | 326 | 11 099NR8 | 099nr8 myocastor c  |
| 45 | 602   | 30.6 | 326 | 11 099NR7 | 099nr7 muscardinus  |

## ALIGNMENTS

|   |              |      |         |  |  |
|---|--------------|------|---------|--|--|
| RESULT 1  |              |      |         |  |  |
| ID 095977   | PRELIMINARY: | PRT: | 384 AA. |  |  |
| AC 095977:  |              |      |         |  |  |
| DT 01-MAY-1999 (TREMBLrel. 10, Created)   |              |      |         |  |  |
| DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)                                    |              |      |         |  |  |
| DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)                                  |              |      |         |  |  |
| DE Putative G-protein coupled receptor, EDG6 precursor (Hypothetical 41.6 kDa protein). |              |      |         |  |  |
| DE  |              |      |         |  |  |
| GN EDG6   |              |      |         |  |  |
| OS Homo sapiens (Human).  |              |      |         |  |  |
| OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;                              |              |      |         |  |  |
| OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.                            |              |      |         |  |  |
| OX NCBI_TaxID=9606;   |              |      |         |  |  |
| RP [1]  |              |      |         |  |  |
| RP SEQUENCE FROM N.A.   |              |      |         |  |  |
| RX MEDLINE-99009331; PubMed-9790765;  |              |      |         |  |  |
| RA Graeler M.H., Bernhardt G., Lipp M.;   |              |      |         |  |  |
| RT "EDG6, a novel G protein-coupled receptor related to receptors for                   |              |      |         |  |  |
| RT bioactive lysophospholipids, is specifically expressed in lymphoid                   |              |      |         |  |  |
| RT tissue.";  |              |      |         |  |  |
| RL Genomics 53:164-169(1998).   |              |      |         |  |  |
| RL [2]  |              |      |         |  |  |
| RP SEQUENCE FROM N.A.   |              |      |         |  |  |
| RC TISSUE=BL00D;  |              |      |         |  |  |
| RA Strausberg R.;   |              |      |         |  |  |
| RL Submitted (0CT-2001) to the EMBL/Genbank/DBJ databases.                              |              |      |         |  |  |
| DR EMBL: AJ000479; AAA04118.1; -  |              |      |         |  |  |
| DR EMBL: BC014970; AAA14970.1; -  |              |      |         |  |  |
| DR InterPro: IPR004064; EDG6receptor.   |              |      |         |  |  |
| DR InterPro: IPR000276; GPCR_Rhodospn.  |              |      |         |  |  |
| DR Pfam: PF00001; 7tm_1; 1.   |              |      |         |  |  |
| DR PRINTS: PR001526; EDG6RECEPTOR.  |              |      |         |  |  |
| DR PRINTS: PR00237; GPCR_RHODOPSIN.   |              |      |         |  |  |
| DR PRINTS: PR01523; SLPRECEPTOR.  |              |      |         |  |  |
| DR PRINTS: PR01523; SLPRECEPTOR.  |              |      |         |  |  |
| DR PROSITE: PS00262; G_PROTEIN_RECPT_F1_2; 1.   |              |      |         |  |  |
| DR G-protein coupled receptor; Hypothetical protein; Receptor; signal.                  |              |      |         |  |  |
| KW SIGNAL   |              |      |         |  |  |
| FT  |              |      |         |  |  |

SEQUENCE 384 AA: 41623 MW: 369A7BC56AB46A47 CRC64;  
Query Match 100.0%; Score 1966; DB 4; Length 384;  
Best Local Similarity 100.0%; Pred. No. 4,7e-151;  
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNTGTPVAPESCOOLAAGHSHRLVLYHNHSGRLAGRGPEDEGIGALRGSLVAASCLV 60  
DB 1 MNTGTPVAPESCOOLAAGHSHRLVLYHNHSGRLAGRGPEDEGIGALRGSLVAASCLV 60  
QY 61 VLENLLVLAITSHMRSRRMYVYCLVNTLSDLLTGAAYLANVLSGATFRILAPAQWFL 120  
DB 61 VLENLLVLAITSHMRSRRMYVYCLVNTLSDLLTGAAYLANVLSGATFRILAPAQWFL 120  
QY 121 REGLETTALAASFSLFTFAGERFATWVRPVAESGATKTSRYVFGICWMLAALLGMLP 180  
DB 121 REGLETTALAASFSLFTFAGERFATWVRPVAESGATKTSRYVFGICWMLAALLGMLP 180  
QY 181 LIGMNCICAFDRCSLLPLYSKRYILFCLVIFAGVLATIMGLYGAIFRLVQASGOKARPP 240  
DB 181 LIGMNCICAFDRCSLLPLYSKRYILFCLVIFAGVLATIMGLYGAIFRLVQASGOKARPP 240  
QY 241 AARRKARLLKTYLMTLLAFVYCWGPLEGILLADYFGSNLMAOEYLRGMDWILALAVLNS 300  
DB 241 AARRKARLLKTYLMTLLAFVYCWGPLEGILLADYFGSNLMAOEYLRGMDWILALAVLNS 300  
QY 301 AVNPITTSFRSREVCAVLSFLCCGCLRLGMRGPDCLARAVEAHSGASTDSSLRPDRS 360  
DB 301 AVNPITTSFRSREVCAVLSFLCCGCLRLGMRGPDCLARAVEAHSGASTDSSLRPDRS 360  
QY 361 FRGSRSLSFRRMRREPLSSISSVRSI 384  
DB 361 FRGSRSLSFRRMRREPLSSISSVRSI 384

RESULT 2  
Q920L1 PRELIMINARY; PRT; 386 AA.  
ID Q920L1  
AC Q920L1;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAR-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE G protein-coupled receptor precursor.  
GN EDG6 OR EDG6.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BAIB/C; TISSUE=FETAL SKIN-DERIVED;  
RX MEDLINE=99009311; PubMed=9790765;  
RA Graeler M.H., Bernhardt G., Lipp M.;  
RT "EDG6, a novel G protein-coupled receptor related to receptors for  
RT bioactive lysophospholipids, is specifically expressed in lymphoid  
RT tissue";  
RL Genomics 53:164-169(1998).  
DR EMBL: A3060674; CA06847.1; -.  
DR MGD: MGI:1333809; Edg6.  
DR InterPro: IPR004064; EDG6receptor.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm.1; 1.  
DR PRINTS: PRO1526; EDG6RECEPTOR.  
DR PRINTS: PRO0237; GPCR\_RHODOPSIN.  
DR PRINTS: PRO1523; S1PRECEPTOR.  
DR PROSITE: PS50262; G\_PROTEIN\_RECPEP\_FL\_2; 1.  
KM Receptor; Signal.  
FT SIGNAL 1 37 POTENTIAL.  
FT CHAIN 38 386 G PROTEIN-COUPLED RECEPTOR EDG6.  
SQ SEQUENCE 386 AA: 42262 MW: 8AB32CB602BA922 CRC64;

Query Match 82.6%; Score 1623; DB 11; Length 386;

Best Local Similarity 82.8%; Pred. No. 2.5e-123;  
Matches 314; Conservative 28; Mismatches 35; Indels 2; Gaps 2;

QY 6 TPVAPESCOOLAAGHSHRLVLYHNHSGRLAGRGPEDEGIGALRGSLVAASCLVLEN 64  
DB 8 TLVTPESCHRLAAGSHSLVLYHNHSGRLASRGSDGGGIGLGRPSVAAGCLVLEN 67  
QY 65 LVLVAITSHMRSRRMYVYCLVNTLSDLLTGAAYLANVLSGATFRILAPAQWFLREG 124  
DB 68 AMVLAIAITMRSRRMYVYCLVNTLSDLLTGLAYVNVLLSGTRTFQSPVIMFLREG 127  
QY 125 LFTALAASFSLFTFAGERFATWVRPVAESGATKTSRYVFGICWMLAALLGMLP 184  
DB 128 LFMALAASFSLFTFAGERFATWVR-VAESGATKTSRYVFGICWMLAALLGMLP 186  
QY 185 NCLCAFDRCSLLPLYSKRYILFCLVIFAGVLATIMGLYGAIFRLVQASGOKARPPAR 244  
DB 187 NCVCAPRCSLLPLYSKRYILFCLVIFAGVLATIMGLYGAIFRLVQASGOKARPPAR 246  
QY 245 KARRLLKTYLMTLLAFVYCWGPLEGILLADYFGSNLMAOEYLRGMDWILALAVLSAVNP 304  
DB 247 KSRLNTVLMILVAFVYCWGPLEGILLADYFGSNLMAOEYLRGMDWILALAVNSAINP 306  
QY 305 ILYFSRSEVCAVLSFLCCGCLRLGMRGPDCLARAVEAHSGASTDSSLRPDRSFGS 364  
DB 307 LIFSRSREVCAVLSFLCCGCLRLGMRGPDCLRLTEAHSGASTDSSLRPDRSFGS 366  
QY 365 RSLSFRRMRREPLSSISSVRS 383  
DB 367 RSLSFRRMRREPLSSISSVRS 385

RESULT 3  
Q9DDK4 PRELIMINARY; PRT; 362 AA.  
ID Q9DDK4  
AC Q9DDK4;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Sphingosine 1-phosphate receptor.  
GN EDG1.  
OS Brachydanio rerio (Zebrafish) (zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=20563813; PubMed=11112429;  
RA Im D.S., Ungar A.R., Lynch K.R.;  
RT "Characterization of a zebrafish (Danio rerio) sphingosine 1-phosphate  
RT receptor expressed in the embryonic brain";  
RL Biochem. Biophys. Res. Commun. 279:139-143(2000).  
DR EMBL: AF31294; AAG43430.1; -.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm.1; 1.  
DR PRINTS: PRO0237; GPCR\_RHODOPSIN.  
DR PROSITE: PS00237; G\_PROTEIN\_RECPEP\_FL\_1; UNKNOWN\_1.  
DR PROSITE: PS50262; G\_PROTEIN\_RECPEP\_FL\_2; 1.  
KM Receptor.  
SQ SEQUENCE 362 AA: 40583 MW: 43A398871DA2D9D4 CRC64;

Query Match 33.3%; Score 655.5; DB 13; Length 362;  
Best Local Similarity 41.0%; Pred. No. 3.6e-45;  
Matches 152; Conservative 56; Mismatches 136; Indels 25; Gaps 9;

QY 24 LVLVLYHNSGRLAGRGPEDEGIGALRGSLVAASCLVLENLVLVAITSHMRSRRMYV 83  
DB 4 LIARIYNTGKF--RKVHKDGPLKADSVVFTIYCCFIILENVLLVLTIMTKFKHKPMY 61  
QY 84 CLVNTTSDLLTGAAYLANVLSGATFRILAPAQWFLREGILFTALAASFSLFTGER 143  
DB 62 FIGNLISDLAGVYVYNTANILLSGANTYKLTPTQWFRFGSMFVALAASVSLTAIAIER 121

[illegible]

|           |                          |
|-----------|--------------------------|
| RESULT 4  |                          |
| Q9DC35    |                          |
| ID Q9DC35 | PRELIMINARY; PRT; 382 AA |

|    |   |
|----|---|
| DT | 01-JUN-2001 (TREMblrel, 17, Created)                                |
| DT | 01-JUN-2001 (TREMblrel, 17, Last sequence update)                   |
| DT | 01-JUN-2001 (TREMblrel, 17, Last annotation update)                 |
| DE | Endothelial differentiation sphingolipid G-protein-coupled receptor |

OS Mus musculus (Mouse);  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus  
OX NCBI\_TaxID=10090;

RC STRAIN=C57BL/6J; TISSUE=LUNG;  
RX MEDLINE=21085660; PubMed=11217651;  
RA Kawai J., Shinaigawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi Y., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
RA Salto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Salto R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Flatschmann W., Gaasterland T., Gissi C., Kling B., Kochiya H.,  
RA Kuell P., Lewis S., Matsuo Y., Nikido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blae J., Boffelli D., Bojtuna N., Carrinci P., de Bonaldo M.F.,  
RA Brownstein M.-J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustinglich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald C., Rodriguez I., Sakamoto N.,  
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,  
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 409:685-690(2001).  
EMBL, AK004591, BAB3393.1; -  
MGI:1096355; E091.  
DR InterPro: IPR002767, GPCR\_Rhodopsn,  
DR Pfam: PF00001, 7tm.1; 1.  
DR PRINTS: PR00237; GPCRHHODPSN.  
DR PROSITE: PS00237; G\_PROTEIN\_RECPT\_F1.1; UNKNOWN\_1.  
DR PROSITE: PS50262; G\_PROTEIN\_RECPT\_F1.2; 1.  
SQ SEQUENCE 382 AA: 42639 MW: 576409A2BD5C82A CRC64;

|                       |                  |                    |            |             |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match           | 33.2%            | Score 653;         | DB 11;     | Length 382; |
| Best Local Similarity | 39.5%;           | Pred. No. 6.1e-45; |            |             |
| Matches 148;          | Conservative 67; | Mismatches 130;    | Indels 30; | Gaps 8;     |

[illegible]

RESULT 5  
Q9R235  
ID Q9R235 PRELIMINARY; PRT; 382 AA

DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE Lysophospholipid receptor B1.  
GN EDG1 OR IPB1.

OC Mus musculus (mouse);  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129SVJ;  
RX MEDLINE=99133320; PubMed=991453;  
RA Zhang G., Cortes J.J., Weiner J.A., Fukushima N., Chun J.;  
RT "Comparative analysis of three murine G-protein coupled receptors  
RL activated by sphingosine-1-phosphate.";  
RL Gene 227:89-99(1999).  
DR EMBL; AF108019; AAD16975.1; -.  
DR MGD; MGI:1096355; Edg1.  
DR GMPD; MGI:1096355; Edg1.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR InterPro: IPR004061; 51Receptor.  
DR Pfam: PF00001; 7tm.1.1; 1.  
DR PRINTS; PR00237; GPCRHOOPS.  
DR PRINTS; PR01523; 51RECEPTOR.  
DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KM Receptor.

|                       |        |                    |        |             |
|-----------------------|--------|--------------------|--------|-------------|
| Query Match           | 33.0%; | Score 649;         | DB 11; | Length 382; |
| Best Local Similarity | 39.5%; | Pred. No. 1.3e-44; |        |             |

[illegible]

```

OY 121 REGLETTAASTSELTETPAEERATAMWRPAEAGARTKSNVYFICLWMLAALLGMLP 180
Db 120 REGSMFVALSSVSTSLIAIETRTYMLKMLHNG-SNSSFELLISACWAWISLITGELP 178
OY 181 LIGMNCICAFPCRCSLLPEYSEKRTIELCLVIFAGVLATIMGELYCAIFRLOVASOQ----- 235
Db 179 SMGMWICISLSSCSVPEPLYKHYILFCTYFTILLSTIYLICRITSLVYRTSRILTFR 238
OY 236 ----KAPPPARRRRARLLKTVLMLLAFLVYCWPJDFGLLLADYFGSNIWAOEYLRGMW 291
Db 239 KNISAAAR--SSEKSLMLKTVIIVLSEVFIACWAPDFITILLIYV-GCKARTCIIUKAEY 295
OY 292 ILALAVLNSAVNPITITSFGRSEVCAVLSFL-CCGCLRIGMRG-----PGDCLARAVE 343
Db 296 FLVLAVLNSGNPIITLYLTNKEMRAIRIYSCCKCPGDSAGKFRPIITPGMEFSR--- 352
OY 344 AHSACSTTDSLRPR 358
Db 353 -----SKSDNSSHPQ 362

```

```

RESULT 6
09H228      PRELIMINARY;      PRT;      398 AA.
AC 09H228;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Sphingosine 1-phosphate receptor Edg-8 (SPPR) [putative G-protein
DE coupled receptor].
GN GPCR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Im D., Lynch K.R.;
RT "Characterization of a human sphingosine 1-phosphate receptor, Edg-
RT 8."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kochapalli R., Loughran T.P. Jr.;
RT "Characterization of a novel human gene (hsppr) with homology to rat
RT sphingosine 1-phosphate receptor and differential expression in LGL
RL leukemia patients."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT "Identification of G protein-coupled receptor genes from the human
RT genome sequence."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF317676; AAC38113.1; -
DR EMBL; AF331840; AAL57041.1; -
DR EMBL; AB083602; BAB89315.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR01561; BDG8RECEPTOR.
DR PRINTS; PR00237; GPCRHDOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE      398 AA: 41775 MW; 34AD6C40D8250BF CRC64;

Query Match      32.8%; Score 645.5; DB 4; Length 398;
Best Local Similarity 40.8%; Pred. No. 2.5e-44;
Matches 154; Conservative 50; Mismatches 126; Indels 47; Gaps
7,
0Y      22 SRLIVHYNNGSLARGPDEDGALRGISVAASCLVYENILVLAATSHMRSRWY 81
I :|||||:::| | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db      12  SEVIVLHNNTYCKLNGARQPGAGIRAAVYCLANCAIVENLAVLIVGRHFRFAPM  71
Qy      82  YVCVYNITLSDLTGCAVLAVNLGSGARTFRLAPAQWELREGLEFTLAASTFSLEFTAG  141
Db      72  FLTLSTSLTSLDMLGAAVAANITLSSGPLTKLSPALMFARFGVFAVLTAHSVLSLAIAL  133
Qy     142  ERFATMVR---PVASEGATKTSRYGTYGTGICWLLAAILGMLPLLGWNCUCADFRCSSL  197
Db     132  ESSLTMARGAPAPVSSRG----RTLLMAAAAAMGVSLTGLLPLTGLGNCCRLDCACTVL  186
Qy     198  PLYSRKRLTFCVIVPAGVATIMTIGYGAIFRLVOASGOKAP-RBA-----ARRAR-  247
Db     187  PLYAAAYVLFVCYIAFVGLLAICALIYARITQVBARMARLIPARCTAGTTSRRRRPRS  246
Qy     248  -RLKTVLMLLAFVLCWGPLFGLLADVFEGSSNMAOETLYRGMWITLALAVLSANVPIL  306
Db     247  IALLRTLSVLLALFAVCGKPLELLLLLDVACPARTCEVYLQA-DPFLGLAMANSLLNPIT  305
Qy     307  YSFRREVCRAVLSFLCCG-----CLRLGMRGPGDCLAR  340
Db     306  YTLTRNDRLRAHLRLVCCGRHSRGDEPSGQSQASAAEAASGGLRCLPGLDGSFSGSER  365
Qy     341  AVEAHSGASTDSSLRP  357
Db     366  SSPQRDGLDTGGSTGSP  382

```

| Query Match               | Best local Similarity                  | Score                     | DB | Length |
|---------------------------|--|---------------------------|----|--------|
| Matches 150; Conservative | 62; Mismatches 125; Indels 19; Gaps 9; | 42.1%; Pred. No. 3.5e-44; |    | 378;   |

|   |  |  |     |
|---|--|--|-----|
| QY  | 28   | HYSSGSLGAR-GGPEDEGGGALRGSLVAAASCLVYLENTLLAATTSNHRSRWYYCYLV   | 86  |
| Db  | 21   | HYDVGKLAGRLRNDPEGGGLITITITFLVTCSEFIENMLVALWKNKPHNNRYFETIG    | 80  |
| QY  | 87   | NTLISDLTGAAYLANVLNLSGARTFFRLAPQWFLREGLEFTLAASTESLFTAGEPAT    | 146 |
| Db  | 81   | NLALCDLLAGIAKVNILMSGRKTFSLPYWFLREGSFVALGASTCSLLAIALEHRLT     | 140 |
| QY  | 147  | MWRPVAESGAKTSRYVGTIGLCWLLAALGLMPLIGMNCCLAFDRCSLLPLYSKRYIL    | 206 |
| Db  | 141  | MIR-MRPYDANKHRYEFLILGMCMCLAFSGALPILGMNCLENFPDCSTIILPLYSKRYIA | 199 |
| QY  | 207  | FCEVITPFGVALITMGIXGAFELRLVOASGOKAPPAPRARRLKTYLMILLAFLVCGNP   | 266 |
| Db  | 200  | FLISIFALLVTVIILARIYLCKSSSSRRVANHNSEK-SMALLRLVYVSVFLACWSP     | 258 |
| QY  | 267  | LEGLLDADVGSNLMADQ--YLKGMWITALLVNSAVNPDIYFSREVCRAVLSFLCC      | 324 |
| Db  | 259  | LIIFLFLIDV---ACRAKECSILFKSQMFIMLATVNSAMNPVITTLAKMKRAFR-LVC   | 314 |
| QY  | 325  | GCLRLKRGQDDCLARAVEAHSASATSDSLRPDSFROGSRSLSPFRMRPELSSISS      | 380 |
| Db  | 315  | GCL---VKRGK-----TQASPMQPALDPS-RSKSSSSNNSSHPVKXEDLPVAVT       | 360 |
| RESULT 8  |  |  |     |
| Q9P008  | 09P008   | PRELIMINARY; PRT; 384 AA.                                    |     |
| AC  | Q9P008   |  |     |
| DT  | 01-MAY-2000 (TEMBLrel. 13, Created)                                    |  |     |
| DT  | 01-MAY-2000 (TEMBLrel. 13, Last sequence update)                       |  |     |
| DT  | 01-OCT-2001 (TEMBLrel. 18, Last annotation update)                     |  |     |
| DE  | EDG-3.   |  |     |
| GN  | EDG-3.   |  |     |
| OS  | Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).               |  |     |
| OC  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |  |     |
| OC  | Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neocelostei;      |  |     |
| OC  | Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;        |  |     |
| CC  | Tetraodontidae; Takifugu.  |  |     |
| CC  | NCBI_Taxid=31033;  |  |     |
| CC  | [1]  |  |     |
| RN  | SEQUENCE FROM N.A.   |  |     |
| RP  | MEDLINE=99439719; PubMed=10508926;                                     |  |     |
| RX  | Yamaguchi F., Yamaguchi K., Tokuda M., Brenner S.;                     |  |     |
| RA  | "Molecular cloning of EDG-3 and N-Shc genes from the puffer fish, Fugu |  |     |
| RT  | rubripes, and conservation of syteny with the human genome."           |  |     |
| RL  | EMBL; AF164114; AAF07896.1; -  |  |     |
| DR  | InterPro: IPR004062; EDG3Receptor.                                     |  |     |
| DR  | InterPro: IPR000276; GPCR_Rhodopsn.                                    |  |     |
| DR  | InterPro: IPR004061; S1PReceptor.                                      |  |     |
| DR  | Pfam; PF00001; 7tm_1; 1.   |  |     |
| DR  | PRINTS; PR01524; EDG3RECEPTOR.   |  |     |
| DR  | PRINTS; PR00237; GPCRHOPOPSN.  |  |     |
| DR  | PRINTS; PR01523; S1PReceptor.  |  |     |
| DR  | PROSITE; PS00237; G_PROTEIN_RECP_FL1; UNKNOWN_1.                       |  |     |
| DR  | PROSITE; PSS0262; G_PROTEIN_RECP_FL2; 1                                |  |     |
| SO  | SEQUENCE 384 AA; 42712 MW; 92BDDDF6BBD0E3BF CRC64;                     |  |     |
| Query Match 32.6%; Score 640; DB 13; Length 384;              |  |  |     |
| Best Local Similarity 41.9%; Pred. No. 6, 8e-44;              |  |  |     |
| Matches 147; Conservative 58; Mismatches 128; Indels 18; Gaps |  |  |     |
| QY  | 24   | LIVLHYNHSGLARCGPEDS--GLGALGLSYAASC-LVYLENLLAATTSNHRSR        | 79  |
| Db  | 6  | LIVLHYNTGKLDHR--PLYGTSPTGRDPTIAFLVAVCSFIILENTLVLAIAWKNNRPHN  | 63  |
| QY  | 80   | WYVYLIVVITSDLTGAAYLANVLNLSGARTFRLAPQWFLREGLEFTLAASTESILPT    | 139 |
| Db  | 64   | RHYFTIGNALCDLLASAVAYLVNLLSGEKTLDQSLPYLWVVRGSGNFVILGASIFELAI  | 123 |
| QY  | 140  | AGERAPAVWRVAESGAKTSRYVGTIGLCWLLAALGLMPLIGMNCCLAFDRCSLLPL     | 199 |

```

Db      124 ALEKRLTMIRK-WRPYDASKNNRVFLITCTWLVANVLLGALILGNCNGLNPDOSTIPL 182
QY      200 YSKRYILECLVIFACVLATIMGLYGAIFRLVQASQKAPPAARRKARLLKTYMILLA 259
Db      183 YTKRYVACICVIFVILLAMSVLYARIYILVKSQQKSHRNSHEAMSLRTVIYGV 242
QY      260 FLVCGPFLFGLLADVFGSSNMAQE----YLKGMWILALAVNSAVNPITYSPSRVC 315
Db      243 FIACMPFLFVLLLDV-----ACRRPCPIYKKAMFLAVAVLNSAMPITYTLASREMR 296
QY      316 RAVLSFLCCGCGRLGMRGPDCLARAVE-AHSGASTTSSLRPDSRFGSR 365
Db      237 RAFLGLVCGVCTIRGNGSGNDGSKNOFQEPSRSRKSMSQTHHPNDSQSSR 347

RESULT 9
O9JKMS
ID      O9JKMS      PRELIMINARY:      PRT:      400 AA.
AC
DB      01-OCT-2000 (TREMBLrel. 15, Created)
DT      01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DN      Sphingosine 1-phosphate receptor Edg-8.
GN      EDG-8.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_Taxid=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20261551; PubMed=10799507;
RA      Im D.-S., Heise C.E., Ancellin N., O'Dowd B.F., Shen G.-J.,
RA      Heavens R.P., Rigby M.R., Hla T., Mandala S., McAllister G.,
RA      George S.R., Lynch K.R.;
RT      "Characterization of a novel sphingosine 1-phosphate receptor, Edg-
RT      8."
RL      J. Biol. Chem. 275:14281-14286(2000).
DR      EMBL, AF233649; AAF35912.1; -
DR      InterPro: IPR002376; GPCR_Rhodopsn.
DR      InterPro: IPR004061; S1PReceptor.
DR      Pfam: PF00001; 7tm_1; 1.
DR      PRINTS: PRO1561; EDG8RECEPTOR.
DR      PRINTS: PRO0237; GPCR_RHODOPSIN.
DR      PRINTS: PRO1523; S1PRECEPTOR.
DR      PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; UNKNOWN_1.
DR      PROSITE: PS50262; G_PROTEIN_RECPT_F1_2; 1.
KM      Receptor.
SQ      SEQUENCE 400 AA; 42368 MW; 71B396DC4DEED0FEB CRC64;

Query Match      32.5%; Score 638; DB 11; Length 400;
Best Local Similarity 44.6%; Pred. No. 1e-43;
Matches 145; Conservative 47; Mismatches 113; Indels 20; Gaps
5,

QY      22 SKLIYLVNHSRLAGRGSPEDGIGALRGLSVAASCLVLENLYLAITSHMSRRVY 81
Db      12 SEVIVLVNHYTKLGARQYQPGAGLRADAAVCAIAVFLENLVLLVLRHPRHFM 71
QY      82 YVCIVNITLSDLLTGAAYLVANVLSGARFRLRPAQWFLREGLTFLAASFTSLFTAG 141
Db      72 FLLTSLTSLDLLAGAAVYNITLSSPLTLLRSLPALMFAREGGVALLAASVLSLALAL 131
QY      142 ERFATVVRVAASGATKTSRVYGFGLICWLLAALGMLPLGMNCLCAFDRCSSLLPYS 201
Db      132 EKHILMAR-GRAPASASRRTLAMVAAMGLSLGLLIPALGMNCGRLRACSTVLPYA 190
QY      202 KRYILFCLVIRGVLATIMGLYGAIFRLVQASQK-APPAAR-----KARL 249
Db      191 KAVVLFCLVAFGLIAALCALYARIYQYVANAARRLRADGSRRTATSSSRKHTPSTAL 250
QY      250 IATVIMLILAFVCGPFLFGLLADVFGSSNMAQETLRGMWILALAVNSAVNPITYSF 309
Db      251 IRLVSLVLAFAVACGPFLLTLLLDV-ACPARRCVYLQADPFLGLAMNSLLNPITYTF 309

```

OY 310 RSREVCRAVLSPCCGCLRLGMRP 334  
 : : : : :  
 Db 310 TNRDLRHALLRLCCG-----RGP 328

## RESULT 10

O91X56 PRELIMINARY; PRT: 400 AA.  
 AC 091X56;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
 DE Similar to sphingosine 1-phosphate receptor.  
 GN EDG8.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Strausberg R.;  
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC012332; AAI1232.1; -  
 DR MGI: 2150641; Edg8.  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PFO0001; 7tm\_1; 1.  
 DR PRINTS: PR01561; EDG8RECEPTOR.  
 DR PROSITE: PS00337; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
 DR PROSITE: PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 400 AA: 42330 MW; AA3118E8515F3F4 CRC64;

Query Match 32.3%; Score 636; DB 11; Length 400;  
 Best Local Similarity 41.9%; Pred. No. 1.5e-43;  
 Matches 153; Conservative 51; Mismatches 125; Indels 36; Gaps 8;

OY 22 SRLVLVHNHSGRLAGRGPEDEGLGALRGLSVAASCLVLENLVLVAITSHMSRRMV 81  
 : : : : :  
 Db 12 SEVIIVHYNYTKLNGARVQPGAGLRADAACVLAFCATVLENLAVLVVHPRFHAPM 71  
 : : : : :  
 OY 82 YVCLVNTISDLTGAAVYANVLSGARTFRLAPQWFLREGSLFTALASTFSLLFTAG 141  
 : : : : :  
 Db 72 FLLSLSTLSDLLAGAAVNTNLTLSGPTLRSLPALMFAREGGVVALAASVLSLALAL 131  
 : : : : :  
 OY 142 EEPATVVRVAESGATKTSRVYGFGLGCLLALALGMLPLGNNCLCAFDCSSLLPYXS 201  
 : : : : :  
 Db 132 EKHLMAR-GRAPASARARTLMAAVAMAGSLGLGLPALGWNCLGRLETCSTVPLPYA 190  
 : : : : :  
 OY 202 KRYILFECVIFAGVLAITMGLGATFRLVQASGQK-APRPAAR-----KARRL 249  
 : : : : :  
 Db 191 KAVVLECVLAFLGILALCALYARLYCOVRANARLRAGPGSRRTSSRSRHTPRSLAL 250  
 : : : : :  
 OY 250 LKTVLMLLAFLVCGPFLGILLADYFGSNLMAQEYLGMDITLAVLNSAVNDIYFSF 309  
 : : : : :  
 Db 251 LRTLSVLLAFAVCGPFLFLDLIDY-ACPARACPVLLQADPFLGLAMANSILNIITYF 309  
 : : : : :  
 OY 310 RSREVCRAVLSPCCGCLRLGMRP-----GDCLARAVEAHSAGS-----TTDSS 354  
 : : : : :  
 Db 310 TNRDLRHALLRLCCG-----RGPCCNDSSNSLORSPSA-AGPSGGGLRRCCLPPTLDRS 362  
 : : : : :  
 OY 355 LRRPD 359  
 : : : : :  
 Db 363 SSPSE 367

## RESULT 11

O9NYN8 PRELIMINARY; PRT: 382 AA.  
 AC 09NYN8; O9BY4;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
 DE G protein-coupled sphingolipid receptor (Hypothetical 42.8 kDa  
 protein) (EDG1).  
 GN CHEDG1 OR EDG1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Tigy1 G.U. Sr., Wang D. Sr.;  
 RT "Correction to Human EDG1 sequence."  
 RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LUNG;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RN (3)  
 RP SEQUENCE OF 41-366 FROM N.A.  
 RX MEDLINE-21082082; Pubmed-11214319;  
 RA Murphy W.J., Elizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,  
 O'Brien S.J.;  
 RT "Molecular phylogenetics and the origins of placental mammals."  
 RL Nature 409:614-618(2001).  
 DR EMBL: AF233365; AAF43420.1; -  
 DR EMBL: BC018650; AAI18650.1; -  
 DR EMBL: AY011725; AAK01993.1; -  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR InterPro: IPR004061; S1PReceptor.  
 DR Pfam: PFO0001; 7tm\_1; 1.  
 DR PRINTS: PR00237; GPCRHOOPSN.  
 DR PRINTS: PR01523; S1PRECEPTOR.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
 DR PROSITE: PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Hypothetical protein; Receptor.  
 SQ SEQUENCE 382 AA: 42810 MW; 0CCB8685A5E1BAD2 CRC64;

Query Match 32.3%; Score 635.5; DB 4; Length 382;  
 Best Local Similarity 40.5%; Pred. No. 1.6e-43;  
 Matches 136; Conservative 64; Mismatches 121; Indels 15; Gaps 6;

OY 1 MNATGPVAPESCCOLAAGSHRLVLYHNHSGRLAGRGPEDEGLGALRGLSVAASCLV 60  
 : : : : :  
 Db 1 MGPTSVPLVAHRSVSDYVNDIYVHNYTKL-NISDKENSKIKLSVVFILLCFI 59  
 : : : : :  
 OY 61 VLENLVLVAITSHMSRRVYVCLVNTISDLTGAAVYANVLSGARTFRLAPQWFL 120  
 : : : : :  
 Db 60 ILENIFVLLTIWTKKFKHPRMYYFIGNLALSDLAGVAYANLTLGATFYKLTLPQWFL 119  
 : : : : :  
 OY 121 REGILFTALASTFSLLFTAGEEPATVVRVAESGATKTSRVYGFGLGCLLALALGMLP 180  
 : : : : :  
 Db 120 REGMEFVALNSVPSLALAIERYITMLKKKLNG-SNNRFLLLISACVVISLILGLP 178  
 : : : : :  
 OY 181 LIGNNCLCAFDRCSLLPLYSKRYILFECVIFAGVLAITMGLGATFRLVQASGQ----- 235  
 : : : : :  
 Db 179 IMGNCISALSSCSSTVPLVHKHYILFECTVFLLLLSIYIVCRISLVTRSRRLTR 238  
 : : : : :  
 OY 236 ----KARPAARKARRLTVMILLIAPLVCGPFLGILLADYFGSNLMAQEYLGMDM 291  
 : : : : :  
 Db 239 KNISKASR--SSEKSLALPTIIVISVFLACWAPFILLIDY-GCKYKTCILIRAEY 295  
 : : : : :  
 OY 292 ILALAVNSAVNPDIYSFRSEVCRAVLSFL-CCGC 326  
 : : : : :  
 Db 296 FLVLAIVNSGNTPIITLTITKMKRRAFIRIMSCCK 331  
 : : : : :

## RESULT 12

O99MN8 PRELIMINARY; PRT: 387 AA.  
 AC 099MN8;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Sphingosine-1-phosphate receptor Lpb4 (Fragment).  
 GN EDG8 OR Lpb4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SWISS WEBSTER/NIH;  
 RA Yang A.H., Zhang G., Chun J.J.M.;  
 RT "Molecular cloning of the mouse sphingosine-1-phosphate receptor gene,  
 RT Lpb4.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF327535; AK15485.1; -  
 DR MGI: 2150641; Edg8.  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS: PRO1561; EDG8RECEPTOR.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECPEP\_FL\_1; UNKNOWN\_1.  
 DR PROSITE: PS50262; G\_PROTEIN\_RECPEP\_FL\_2; 1.  
 KW Receptor.  
 FT NON TER  
 SQ SEQUENCE 387 AA; 40953 MW; B56B47672BF5E977 CRC64;  
 Query Match 32.1%; Score 632; DB 11; Length 387;  
 Best Local Similarity 41.9%; Pred. No. 3,1e-43;  
 Matches 152; Conservative 51; Mismatches 124; Indels 36; Gaps 8;  
 QY 24 LYLVAHNSGRLAGRGPEDEGGIGALRGLSVAASGLVLENLVLAATSHMSRRMY 83  
 Db 1 VYLVAHNYTKLKGARYGAGLRADAVCLAVCAFYLENLVLAIVGHRPHAM 60  
 QY 84 CLYNITISDLTGCAATLANVLSGARTFLAPADQFLREGILFTALASTFSLFTAGER 143  
 Db 61 LLSGLTSLDLAGAAVATNILLSGPLTLRLSPALMFAREGGVFVALAASVLSLALALR 120  
 QY 144 FATVPRVAESGATKTSRYVGFIGLCMLAALGMLPLTGNNCCAPDRCSLLPYSKR 203  
 Db 121 HLTMAR-GRAPASARFLTAMAVAMGASLLGLRLPALGNNCCGRLETCSTYLPIYANA 179  
 QY 204 YILFCVITFAGVATITMGITGATFRLVQASGOK-APRPAR- 251  
 Db 180 YLFCVIAFGILTAICALYARIYCOVRANARLRAGGSRATSSRSRHTPRSLALR 239  
 QY 252 TVIMTILAFVLCWGPFLGLLADVFSGNLMAGQYLKGMWILALAVNSAVNPITYSFS 311  
 Db 240 TLSVLLAFVACWGPFLFLLLDDV-ACPARACPVLLQADPFLGLAMANSLLNPITYTFN 298  
 QY 312 RENCRAVLSFLCCGCLRLGMGP-----GDCIARAVEHSGAS-----TTPSSUR 356  
 Db 299 RDLNHALRLRLCCG-----RQPCNODSSNSLQNSPSA-AGPSGGGLRCLPPLDPRSSS 351  
 QY 357 PRD 359  
 Db 352 PSE 354  
 RESULT 13  
 ID 090Y79 PRELIMINARY; PRT; 400 AA.  
 AC 090Y79;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Growth factor-regulated G protein-coupled receptor Nrg-1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=94453967; PubMed=10532805;  
 RA Glickman M., Malek R.L., Wittek-Black A.E., Jacob H.J., Lee N.H.;  
 RT "Molecular cloning, tissue-specific expression, and chromosomal  
 RT localization of a novel nerve growth factor-regulated G-protein-  
 RT coupled receptor, nrg-1.";  
 RL Mol. Cell. Neurosci. 14:141-152(1999).  
 DR EMBL: AF15249; AAF15395.1; -  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR InterPro: IPR004061; S1PReceptor.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS: PRO1561; EDG8RECEPTOR.  
 DR PRINTS: PRO0237; GPCR\_RHODOPSIN.  
 DR PRINTS: PRO1523; S1PRECEPTOR.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECPEP\_FL\_1; UNKNOWN\_1.  
 DR PROSITE: PS50262; G\_PROTEIN\_RECPEP\_FL\_2; 1.  
 KW Receptor.  
 FT NON TER  
 SQ SEQUENCE 400 AA; 42368 MW; 72B46CCB3BD24D4B CRC64;  
 Query Match 31.9%; Score 628; DB 11; Length 400;  
 Best Local Similarity 44.0%; Pred. No. 6,7e-43;  
 Matches 143; Conservative 47; Mismatches 115; Indels 20; Gaps 5;  
 QY 22 SRLIVAHNSGRLAGRGPEDEGGIGALRGLSVAASGLVLENLVLAATSHMSRRMY 81  
 Db 12 SEVILAHNYTKLKGARYGAGLRADAVCLAVCAFYLENLVLAIVGHRPHAM 71  
 QY 82 YLVLNITISDLTGCAATLANVLSGARTFLAPADQFLREGILFTALASTFSLFTAG 141  
 Db 72 FLISGLTSLDLAGAAVATNILLSGPLTLRLSPALMFAREGGVFVALAASVLSLALAI 131  
 QY 142 ERPAVPRVAESGATKTSRYVGFIGLCMLAALGMLPLTGNNCCAPDRCSLLPYTS 201  
 Db 132 ERLITMAR-GRAPASARFLTAMAVAMGASLLGLRLPALGNNCCGRLETCSTYLPIYA 190  
 QY 202 KRTILFCVITFAGVATITMGITGATFRLVQASGOK-APRPAR- 249  
 Db 191 KAVYLFCVIAFGILTAICALYARIYCOVRANARLRAGGSRATSSRSRHTPRSLAL 250  
 QY 250 LKVLNITILAFVLCWGPFLGLLADVFSGNLMAGQYLKGMWILALAVNSAVNPITYSF 309  
 Db 251 LRLSVLLAFVACWGPFLFLLLDDV-ACPARACPVLLQADPFLGLAMANSLLNPITYTF 309  
 QY 310 RSREVCRAVLSFLCCGCLRLGMGP 334  
 Db 310 TNRDLNHALRLRLCCG-----RGP 328  
 RESULT 14  
 ID 099NR2 PRELIMINARY; PRT; 326 AA.  
 AC 099NR2;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)  
 DE EDG1 (Fragment).  
 OS Dipodomys heermanni (Kangaroo rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Heteromyidae;  
 OC Dipodomys; Dipodomys.  
 OX NCBI\_TaxID=10018;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21082082; PubMed=11214319;  
 RA Murphy W.J., Bizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,  
 RA O'Brien S.J.;  
 RT "Molecular phylogenetics and the origins of placental mammals.";  
 RT Nature 409:614-618(2001).  
 RL EMBL: AY011710; AA01978.1; -  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS: PRO0237; GPCR\_RHODOPSIN.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECPEP\_FL\_1; UNKNOWN\_1.

